

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:27:07 ; Search time 96.19 Seconds
(without alignments)
8107.378 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtgacccacgctgcg.....aaaaaaaaaaaaaaaa 3117

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	68.4	2.2	75	1 T99072	Alpha epithelial s
2	59	1.9	231	1 T21330	Human gene signatu
3	56.2	1.8	795	1 V55830	FLGA insert stabil
4	52.2	1.7	799	1 V55831	Nucleotide sequenc
5	52.2	1.7	9600	1 V21683	Vector plasmid PCM
6	52.2	1.7	10596	1 Q51731	Plasmid pCisEBON f
7	52.2	1.7	10596	1 T40348	Plasmid pCisEBON f
8	52.2	1.7	10596	1 X15650	Nucleotide sequenc
9	50	1.6	1620	1 V60840	Partial human acid
10	48.2	1.5	558	1 T64563	Mouse thymus and a
11	46.8	1.5	1525	1 V22704	Mouse recombinase
12	45.4	1.5	1129	1 X27340	Human secreted pro
13	45	1.4	45	1 T99077	Probe Ena-A7 for a
14	45	1.4	45	1 T99078	Probe Ena-A9 for a
15	44.6	1.4	1096	1 X22275	Human secreted pro
16	44.6	1.4	1631	1 T22795	Beta-1 integrin mo
17	44.6	1.4	6644	1 X33181	Base sequence of t
18	44.6	1.4	7372	1 X33182	Base sequence of t
19	44.6	1.4	7797	1 X33180	Cowpox virus bsr f
20	44.6	1.4	7996	1 X33184	Base sequence of t
21	44.4	1.4	960	1 X28162	Rat Acid sensitive
22	44.4	1.4	3562	1 V60839	Rat acid sensing i
23	44.2	1.4	1632	1 V68059	Neurodegenerative
24	44.2	1.4	1977	1 X27383	Human secreted pro
25	44.2	1.4	2711	1 V68056	Neurodegenerative
26	44.2	1.4	2955	1 V68057	Neurodegenerative
27	44	1.4	1348	1 X24749	Human interleukin-
28	43.8	1.4	2823	1 T35233	Natural killer lyt
29	43.8	1.4	3350	1 T18200	Infectious bursal
30	43.8	1.4	5503	1 V18187	Fanconi anaemia of
31	43.6	1.4	1174	1 T39050	cDNA encoding cell
32	43.6	1.4	1174	1 V39096	Monocomponent endo
33	43.6	1.4	1537	1 V59804	Human secreted pro

34 43.4 1.4 437 1 X22245 Human secreted pro
35 43.4 1.4 1338 1 Q65607 Rabbit zona pelluc
36 43.4 1.4 1338 1 V64789 Rabbit 2PC cDNA. I
37 43.4 1.4 1829 1 V26298 Phosphate starvati
38 43.4 1.4 2132 1 X33816 Coding sequence fo
39 43.4 1.4 2549 1 N91467 Sequence of human
40 43.4 1.4 3557 1 N60801 Human pro-growth h
41 43.4 1.4 3784 1 Q05326 Sequence encoding
42 43.4 1.4 5018 1 N70646 Clone 23B6p102 enc
43 43.4 1.4 5018 1 N50150 Sequence of the cD
44 43.2 1.4 1355 1 N80457 Sequence of cDNA e
45 43.2 1.4 2962 1 X28161 Rat Acid sensitive

ALIGNMENTS

RESULT 1

T99072 ID T99072 standard; DNA; 75 BP.
AC T99072:
24-MAR-1998 (first entry)
DE Alpha epithelial sodium channel a coding sequence.
KW Alpha epithelial sodium channel; alphaENACA; alphaENACb; binding assay;
KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
KW salt substitute; salty taste blocker; ss.
OS Rattus rattus.
PN US5693756-A.
PD 02-DEC-1997.
PF 23-JAN-1995; 376362.
PR 23-JAN-1995; US-376362.
PR 28-FEB-1994; US-202654.
PA (UYJO) UNIV JOHNS HOPKINS.
PI Blackshaw S, Li X, Snyder SH;
DR WPI; 98-031814/03.
DR P-PSDB; W34529.
PT Alternatively spliced epithelial sodium channel alpha subunit
PT proteins - useful in screening assays for salty taste enhancers or
PT blockers
PS Disclosure; Column 21-22; 33pp; English.
CC This sequence represents the coding sequence for the alpha epithelial
CC sodium channel a (alphaENACA). AlphaENACA and alphaENACb (see W34530)
CC represent the sequences of the invention. The two sodium channels are
CC alternatively spliced forms of the amiloride-sensitive salt channel alpha
CC subunit and can be used in membrane-transport or binding assays to
CC identify substances that enhance or block perception of a salty taste.
CC Enhancers could be used as salt substitutes and blockers could be used to
CC mask salty tastes in foods and pharmaceuticals.
SQ Sequence 75 BP; 11 A; 22 C; 22 G; 20 T;

Query Match 2.2%; Score 68.4; DB 1; Length 75;
Best Local Similarity 99.6%; Pred. No. 4.46e-07;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1545 gggcgcttccttgagacagcctggctgttttccaaagtgcggaagccttagtgt 1604
|||||
Db 6 GGGCGCCTTCCTCGGACGCCCTGGGCTGTCTCTCCAAAGTGTGGAGAGCCTTAGTGT 65
|||||

QY 1605 gatcaactac 1614
|||||
Db 66 GATCAACTAC 75
|||||

RESULT 2

T21330 ID T21330 standard; cDNA to mRNA; 231 BP.

AC T21330; 1996 (first entry)
DT 16-AUG-1996
DE Human gene signature HUMGS02692.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.

SQ Sequence 10596 BP; 2627 A; 2571 C; 3023 G; 2375 T;
 Query Match 1.7%; Score 52.2; DB 1; Length 10596;
 Best Local Similarity 50.9%; Pred. No. 0.013;
 Matches 149; Conservative 0; Mismatches 143; Indels 1; Gaps 1;

QY	1881	cttcgaactcctggtcatcacacttctatctgtacgcgggttcaggagcccgta	1940
DB	2961	CCTCCTGACCCCGCCTCCACCTCCTGCTCTGCCCTCCTGTCTTGCCTGCTTC	2902
QY	1941	gtctccaaggacaggggccagggggtgcaggagaagtggctctcaactcagat	2000
DB	2901	G-CRCCCTGCCCTCCTTGCCCCCTCTGCTCCCTGCCCTCCTGCCCTCCTG	2843
QY	2001	cccgctccggtttctgtctcaactacataccccaccaccttcttttgcccag	2060
DB	2842	TCCCTGCCCTCCTGCTCTGCTGCCCTCCTTGCCCCCTCCTGCTGCTGCC	2783
QY	2061	gaccctctccctggccctgcagagccctccaccctgcctatctactcttag	2120
DB	2782	TCCCTCCTGCTCTGCCCTCCTTGCCCCCTCCTGCTGCCCTCCTGCTGCC	2723
QY	2121	ccctccactggactctggggcgccctgaactgttctgctgtgacctggcggc	2173
DB	2722	TGCCCTCCTGCCCTCCTGCTCTGCCCTCCTGCTGCCCTCCTGCTGCC	2670

RESULT 8
 X15650/C
 ID X15650 standard; DNA; 10596 BP.
 AC X15650;
 DT 10-MAY-1999 (first entry)
 DE Nucleotide sequence of plasmid pCIS.EBON.
 KW Plasmid pCIS.EBON; hepatocyte growth factor; HGF; variant;
 KM HGF receptor activation; ss.
 OS Synthetic.
 PN US5879910-A.
 PD 09-MAR-1999.
 PF 09-FEB-1994; 194087.
 PR 09-FEB-1994; US-194087.
 PA (GENE) GENENTECH INC.
 PI Godowski PJ, Lökke NA, Mark MR;
 DR WPI; 99-203949/17.
 PT New hepatocyte growth factor variants - useful for studying
 PT structure-function relationships in the wild type molecule and for
 PT treating conditions associated with chronic hepatocyte growth factor
 PT receptor activation
 PS Example 1; Fig 6A-E; 40pp; English.
 CC The present sequence represents the nucleotide sequence of pCIS.EBON,
 CC which is used in the course of the invention. The specification
 CC describes a hepatocyte growth factor (HGF) variant (HGFP) comprising an
 CC amino acid (aa) alteration at or adjacent to position 652 of the
 CC wild-type human HGF (huHGF) aa sequence. HGFP may be used in studies to
 CC identify the structure-activity relationships of HGF in order to identify
 CC the functionally important domains in the aa sequence. It may also be
 CC used to identify aa residues which are responsible for the interaction of
 CC HGF with its receptor, and those aa that are responsible for the
 CC biological activity of it. Variants of HGFP which have enhanced receptor
 CC binding affinity (compared to wild-type huHGF) and are more biologically
 CC active than wild type huHGF, may be used as huHGF agonists. Conversely,
 CC variants of HGFP which have enhanced receptor binding affinity (compared
 CC to wild-type huHGF) but which are biologically inactive may be used as
 CC huHGF antagonists, and may be used to block the binding of wild-type
 CC huHGF to its receptor. This permits the treatment of pathogenic
 CC conditions associated with the activation of an HGF receptor, such as
 CC malignancies associated with chronic HGF receptor activation.
 CC HGF-immunoglobulin chimeras may be produced (by standard methods) and
 CC used in protein A purification, immunohistochemistry, and
 CC immunoprecipitation techniques (in place of anti-HGF antibodies) or in
 CC screening studies to identify inhibitors of HGF-HGF interactions.
 CC Sequence 10596 BP; 2627 A; 2571 C; 3023 G; 2375 T;

	Query Match	1.7%; Score 52.2; DB 1; Length 10596;
	Best Local Similarity	50.9%; Pred. No. 0.013;
	Matches 149; Conservative	0; Mismatches 143; Indels 1; Gaps 1;
Qy	1881	cttgagactcctgtgcatcacacttctcatgtctacgccggttcccgaggaccggtactg 1940
Dd	2961	CCTGCTCAGCCGGCGTCCACACTCTGTCTGCTGCCCCCTGCTGTCCTGCTGCCCTCCTCCT 2902
Qy	1941	gtctccaggacgagggccagggtgcagagaggtggcctccactccagcttctctcctt 2000
Dd	2901	G-CTCTGCCCTCTGCTGCCCTCTGCTGCCTGCTGCCCTCTGCTGCCCTCTGCTGCCCTC 2843
Qy	2001	ccccgtccctgttctgtctcaacctacatccccaccacttcttggccccagcagggat 2060
Dd	2842	TCTTGCCCTCTCTGCTGCTGCCCTCTGCTGCCCTCTGCTGCCCTCTGCTGCCCTC 2783
Qy	2061	gaacctcccctggccctgaagccctccaccctgctctactctagttagccccaagtgc 2120
Dd	2782	TCCCTCTGCTGCCCCCTGCTGCCCTCTGCTGCCCTCTGCTGCCCTCTGCTGCCCTC 2723
Qy	2121	ccctccactgacttcggcgccctgacttgttgcctgtgccccggggcgc 2173
Dd	2722	TGCCCTCTGCCCCTCTGCTGCTGCCCTCTGCTGCCCTCTGCTGCCCTCTGCTGTC 2670
RESULT	9	
V60840	ID	V60840 standard; DNA; 1620 BP.
AC	V60840;	
DT	02-FEB-1999	(first entry)
DE	PARTIAL human acid sensing ionic channel gene.	
KW	Human; neuronal cationic channel; amiloride; proton; ASIC; brain; probe;	
KW	acid sensing ionic channel; hybridisation; primer; PCR; amplification;	
KW	modulator; acidity; nociception; pain; taste; inflammation; ischaemia;	
KW	tumour; cerebral neurodegeneration; transgenic animal; knockout animal;	
KW	gene therapy; Alzheimer's; Parkinson's; Huntington's; disease;	
KW	amyotrophic lateral sclerosis; cerebellar ataxia; ds.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	1..1545
FT	/tag= a	
FT	/product= "partial ASIC"	
FT	/note= "acid sensing ionic channel; no start codon	
FT	is given at the 5' end of the sequence"	
PN	W09835034-A1.	
PD	13-AUG-1998.	
PF	11-FEB-1998; F00270.	
PR	28-JUL-1997; FR-009587.	
PR	11-FEB-1997; FR-001574.	
PA	(CNRS) CNRS CENT NAT RECH SCI.	
PI	Bassilana F, Champigny G, Heurteaux C, Lazdunski M,	
PI	Waldmann R, Lingueglia E;	
DR	WPI; 98-447231/38.	
DR	p-PSDB; W68505.	
PT	Protein comprising proton-sensitive neuronal channel - useful for	
PT	screening for analgesics and for treating neurodegeneration	
PS	Claim 13, Page 28-30; 64pp; French.	
CC	This sequence represents a partial gene encoding a human neuronal	
CC	cationic channel that is sensitive to amiloride and activated by protons	
CC	and is designated Acid Sensing Ionic Channel (ASIC). The protein can be	
CC	used to screen for modulators of these channels, particularly to identify	
CC	compounds that modulate perception of acidity, as regards nociception	
CC	(pain) rather than taste. These compounds are used to treat or prevent	
CC	pain associated with acidity (e.g. in cases of inflammation, ischaemia	
CC	or some tumours) and as inhibitors of neurodegeneration caused by	
CC	overexpression of the channels. Antibodies to the protein are used to	
CC	detect the channels in tissues, and to act therapeutically as channel	
CC	modulators. The nucleic acid can be used to generate transgenic,	
CC	particularly knockout, animals for studying ASIC-related disorders,	
CC	also for gene therapy. The channel protein, or its (ant)agonists, can	
CC	be used to treat or prevent cerebral neurodegenerative conditions (e.g.	
CC	Alzheimer's, Parkinson's or Huntington's diseases, amyotrophic lateral	

CC alphaENACb (see W34530) represent the sequences of the invention. The two
CC sodium channels are alternatively spliced forms of the
CC aniloride-sensitive salt channel alpha subunit and can be used in
CC membrane-transport or binding assays to identify substances that enhance
CC or block perception of a salty taste. Enhancers could be used as salt
CC substitutes and blockers could be used to mask salty tastes in foods and
CC pharmaceuticals.
SQ Sequence 45 BP; 11 A; 11 C; 10 G; 13 T;

Query Match 1.4%; Score 45; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1498 gactaccgaagcagagctctgggctattgctattataaactg 1542

Db 45 GACTACCGAAGCAGAGCTCTGGGCTATTGCTATTATAAATG 1

RESULT 15

ID X22275 standard; DNA; 1096 BP.

AC X22275; 18-MAY-1999 (first entry)

DE Human secreted protein gene 20 clone HGCAB62.

KW Human; secreted protein; gene therapy; protein therapy; cancer; weight;

KW tumour; chromosome mapping; forensic; haematological disease; allergy;

KW inflammation; cell proliferation; viral infection; wound healing;

KW modulation; appetite; behaviour; food additive; preservative; ss.

OS Homo sapiens.

PN WO9903990-A1.

PD 28-JAN-1999.

PF 15-JUL-1998; U14613.

PR 18-AUG-1997; US-056361.

PR 16-JUL-1997; US-052661.

PR 16-JUL-1997; US-052870.

PR 16-JUL-1997; US-052871.

PR 16-JUL-1997; US-052872.

PR 16-JUL-1997; US-052873.

PR 16-JUL-1997; US-052874.

PR 16-JUL-1997; US-052875.

PR 22-JUL-1997; US-053440.

PR 22-JUL-1997; US-053441.

PR 22-JUL-1997; US-053442.

PR 18-AUG-1997; US-055683.

PR 18-AUG-1997; US-055724.

PR 18-AUG-1997; US-055725.

PR 18-AUG-1997; US-055726.

PR 18-AUG-1997; US-055946.

PR 18-AUG-1997; US-055952.

PR 18-AUG-1997; US-055985.

PR 18-AUG-1997; US-055989.

PR 18-AUG-1997; US-056359.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Duan R, Feng P, Ferrie AM, Florence KA, Fouad J,

PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PE,

PI Yu G.

DR WPI: 99-132234/11.

DR P-FSDB; Y01447.

PT New nucleic acids encoding secreted human proteins - potentially

PT useful for treating and diagnosing diseases and identifying specific

PT binding agents

PS Claim 4; Page 201-202; 251pp; English.

CC The invention relates to nucleic acid sequences (X22211 to X22282)

CC encoding human secreted proteins (Y01383 to Y01454). The secreted protein

CC gene sequences are deposited with the ATCC under deposit number ATCC

CC 209138, 209139 or 209141. Host cells containing vectors comprising the

CC nucleic acid sequences are used for the recombinant expression of the

CC secreted proteins. The polynucleotide and amino acid sequences are useful

CC for preventing, treating or ameliorating medical conditions e.g. by

CC protein or gene therapy. Pathological conditions can be also diagnosed by

CC determining the amount of the new polypeptides in a sample or by the

CC presence of mutations in the new polynucleotides. The nucleic acid

CC sequences, or its fragments, are useful for chromosome identification and
CC mapping; as antisense and triplex-forming therapeutics; in gene therapy;
CC for (forensic) identification of individuals; as molecular weight
CC markers; to identify related sequences or specific mRNA; in preparation
CC of oligomers and to raise anti-DNA antibodies. Antibodies are useful as
CC immunoassay reagents (including for in vivo imaging) and therapeutically
CC to inhibit or activate particular polypeptides. A very wide range of
CC disorders may be treated with the polynucleotide and polypeptide
CC sequences, e.g. autoimmune or haematological diseases, allergy,
CC inflammation, cancer or other forms of cell proliferation, viral or other
CC infections. The sequences may also be useful in wound healing, to
CC modulate differentiation of embryonic stem cells, to modulate weight,
CC appetite, behaviour etc. and as food additive or preservative. The
CC present sequence represents a gene encoding a human secreted protein
CC (see descriptor line for gene number and clone identification).
SQ Sequence 1096 BP; 191 A; 367 C; 303 G; 231 T;

Query Match 1.4%; Score 44.6; DB 1; Length 1096;
Best Local Similarity 66.3%; Pred. No. 0.32;
Matches 59; Conservative 3; Mismatches 27; Indels 0; Gaps 0;

QY 3029 ccagctctcccaacttgggacgggtaggcaggtactcaataaaggctgttccatcaaaaa 3088

Db 994 CCCAAGGACCCCAATTCTGCTCTCTTCTGGGCAATAATGAGTGTCTCTTTTCAGCARMWA 1053

QY 3089 aaaaaaiaaaaaiaaaaaiaaaaaiaaaaaiaaaaa 3117

Db 1054 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1082

Search completed: September 10, 2000, 23:03:35
Job time: 5788 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:24:57 ; Search time 1087.86 Seconds
(without alignments)
12635.297 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcgaccacgcgtcg.....aaaaaaaaaaaaaaaaaaaa 3117

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues 10495684
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	EST:*
	1: em_est1:*
	2: em_est2:*
	3: em_est3:*
	4: em_est4:*
	5: em_est5:*
	6: em_est6:*
	7: em_est7:*
	8: em_est8:*
	9: em_est9:*
	10: em_est10:*
	11: em_est11:*
	12: em_est12:*
	13: em_est13:*
	14: em_est14:*
	15: em_est15:*
	16: em_est16:*
	17: em_est17:*
	18: em_est18:*
	19: em_est19:*
	20: gb_est1:*
	21: gb_est2:*
	22: gb_est3:*
	23: gb_est4:*
	24: gb_est5:*
	25: gb_est6:*
	26: gb_est7:*
	27: gb_est8:*
	28: gb_est9:*
	29: gb_est10:*
	30: gb_est11:*
	31: gb_est12:*
	32: gb_est13:*
	33: gb_est14:*
	34: gb_est15:*
	35: gb_est16:*
	36: gb_est17:*
	37: gb_est18:*
	38: gb_est19:*
	39: gb_est20:*
	40: gb_est21:*
	41: gb_est22:*
	42: gb_est23:*
	43: gb_est24:*
	44: gb_est25:*
	45: gb_est26:*
	46: gb_est27:*
	47: gb_est28:*
	48: gb_est29:*
	49: gb_est30:*
	50: gb_est31:*
	51: gb_est32:*
	52: em_est20:*
	53: em_est21:*
	54: em_est22:*
	55: em_est23:*
	56: em_est24:*
	57: em_est25:*
	58: em_est26:*
	59: gb_est33:*
	60: gb_est34:*
	61: gb_est35:*
	62: gb_est36:*
	63: gb_est37:*
	64: gb_est38:*
	65: em_est27:*
	66: em_est28:*
	67: em_est29:*
	68: em_est30:*
	69: gb_est39:*
	70: gb_est40:*
	71: gb_est41:*
	72: gb_est42:*
	73: gb_est43:*
	74: gb_est44:*
	75: em_est31:*
	76: em_est32:*
	77: em_est33:*
	78: em_est34:*
	79: gb_est45:*
	80: gb_est46:*
	81: gb_est47:*
	82: em_est35:*
	83: em_est36:*
	84: em_est37:*
	85: gb_est48:*
	86: gb_est49:*
	87: gb_est50:*
	88: gb_est51:*
	89: gb_est52:*
	90: gb_est53:*
	91: gb_est54:*
	92: gb_est55:*
	93: gb_gss1:*
	94: gb_gss2:*
	95: gb_gss3:*
	96: gb_gss4:*
	97: em_gss1:*
	98: em_gss2:*
	99: em_gss3:*
	100: em_gss4:*
	101: gb_gss5:*
	102: gb_gss6:*
	103: gb_gss7:*
	104: gb_gss8:*
	105: gb_gss9:*
	106: em_gss5:*
	107: em_gss6:*
	108: em_gss7:*
	109: em_gss8:*
	110: em_gss9:*
	111: em_gss10:*
	112: em_gss11:*
	113: gb_gss10:*
	114: gb_gss11:*
	115: em_gss12:*
	116: gb_gss12:*

Db	132	TGACAGAACTCCAACTCTGGATGCTTCCATGCTGGAACTCAACACGGTC	191
Qy	1149	cctgcaactgcgcacagagcagaatgaattcatccctcctgctccacagtgcaggggc	1208
Db	192	CCTGATGCTCGCGCGCAGANAGATGACTTCATCCCTGCTGTCCACAGTGACTGGGC	251
Qy	1209	cagggtgatggtgcattggtcaggatgagcctgcctttatggatgatggtggtctcaact	1268

[illegible]

532	CGATTTATGGCGACGTGCACCCAGAGGATGGCCAGTGTGATGTCCTCTGTGTGAGAACCTTTTACCCCTTC	431
D5		
Qy	1389 caagtatacacagcagggtgtgcattcactcctgcttccaggagagaacatgatcaagaagtg 1448	

Accession	Gene	Protein	Accession	Gene	Protein
Db	432	AAAGTACACACAGCGGTGTGTTACCTCCCTGCTTCCAGGAGACGATGATCAAGGAGTG	491		
Qy	1449	tggcgtgcctacatctctacacctaaagcccaaggaggtgagttctctgactaccgaaa	1508		

[illegible]

571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

QY 1629 ctactcag-gtggccatctgtgaag-tcccaggattggatcttcgagatgct-gtccctt 1685

Db 672 GTACTCACGATTGGCCCTCGGTGAANATNCAGGAATCCAGGTCTTCANATGCTATTGCCGA 731

Qy 1686 gcagaacaattacactattacaacaaaaagaa 1717

DB 732 NCGGAACAATTNAAACCGTCAACAACAGNGNAA 763

RESULT 3

AL036530	AL036530	554 bp	mRNA	EST	29-FEB-2000
LOCUS	AL036530	554 bp	mRNA	EST	29-FEB-2000

DEFINITION DXFZp56411062_r1 564 (synonym: hfr2) Homo sapiens cDNA clone DXFZp56411062 5', mRNA sequence.

ACCESSION AL036530
VERSION AL036530.1 GI:5406075
KEYWORDS EST

ORGANISM	human.
SOURCE	EST. NETWORKS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	AUTHORS
1 (bases 1 to 554)	Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wilmann, J.

TITLE	Est (Blum, et al.)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Blum

CONTACT: BRUM H
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by LMD (Ludwig Maximilians University,

Genome Project.
 Munich/Germany) within the cDNA sequencing consortium of the German Research Network for the Genome Project.
 No cDNA sequence available

This clone (DKFZp564I1062) is available at the RZPD in Berlin. NO SI sequence available. Please contact the RZPD, Ressourcenzentrum, Heißenberg 6, 14059

For contact and more information, please contact the author: Rüdiger Rosenbaum, Neurobiology of Learning and Memory, Department of Psychology, University of Hamburg, 20146 Hamburg, Germany; Email: r.rosenbaum@psych.uni-hamburg.de.

```
FEATURES
  source
    Location/Qualifiers
      1. .554
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="DKFZp56411062"
        /clone_lib="564 (synonym: hfbr2)"
        /tissue_type="brain"
        /dev_stage="fetal"
        /lab_host="xl-2blue"
        /note="vector: pAMP1; Site_1: NotI; Site_2: SalI"
BASE COUNT 122 a 154 c 141 g 136 t 1 others
ORIGIN
  Query Match 13.0%; Score 404.8; DB 47; Length 554;
  Best Local Similarity 83.2%; Pred. No. 1.3e-90;
  Matches 460; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
  Qy 1358 gcatgtcccggtcaagaacctttacccttccaagtatacacagcaggtgtgcattcact 1417
  Db 2 GGGATGTTCTCTGAGAACCTTTACCTTCAAGTACACAGCAGGTGTATTCACT 61
  Qy 1418 cctgttccaggagaaacatgatcaagaagtgtggtgtgctgtacatatcttaccctaaagc 1477
  Db 62 CCTGCTTCCAGGAGAGCATGATCAAGGAGTGTGGCTGTGCCTTACATCTTATCCGCGGC 121
  Qy 1478 ccaaggagtgagttctgtgaactaccgaagaagacagagctcctggggtattgtctattata 1537
  Db 122 CCCAAGACGTGGAGTACTGTGACTACAGAAAGCACAGTTCCTGGGGGTACTGCTACTATA 181
  Qy 1538 aactgcagggccttctccttggagacgcctgggtgtgttctcccaagtgtcgaagcctt 1597
  Db 182 ACCTCCAGGTTGACATCTCTCCAGACACCCTGGGTGTTCACCAAGTCCGGAAGCCAT 241
  Qy 1598 gtatgtgatcaactacaactctctgcggctactcaagtggtggtcattgtgaagtcc 1657
  Db 242 CGAGCGGGACCACTGCCAGCTCTCTGTGTGTACTCAAGTGGCCCTCGGTGACATCCC 301
  Qy 1658 aggatgtatcttcagatgctgtcctgtcagaaacattacactattacaacaaagaa 1717
  Db 302 AGGATGGGTCTCCAGATGCTATCGGCACAGAACAAATACACCTCAACACAAAGAGA 361
  Qy 1718 acggagtgtcaaaagctcaactcttcttcaaggagctgaactataaaaactaatcggagt 1777
  Db 362 ATGGAGTGGCCAAAGTCAACATCTTCTCAAGGAGCTGAACATACAAACCAATTCGAGT 421
  Qy 1778 ctctctgtcagatggtgtcagctcctctgtccaaactgggagcagtggaagcctgtggt 1837
  Db 422 CTCCTCTGTACGATGGTTCACCTTCTGTCCAACTGGGGCAGCAGTGGAGCCTGTGCT 481
  Qy 1838 ttggctgtccgtctctctgtgtggagatggcgaagtcattctcgaacctcctgggtca 1897
  Db 482 TCGGCTCTCGGTGTGTCTGTGTGGAGATGGCTGAGCTGCTTTGACCTGCTGGTCA 541
  Qy 1898 tcacattctcat 1910
  Db 542 TCATGTTCTCAT 554
  RESULT 4
  AA925917/c
  LOCUS AA925917 387 bp mRNA EST 04-JUL-1999
  DEFINITION UI-R-A1-es-g-03-0-UI.s1 UI-R-A1 Rattus norvegicus cDNA clone
    UI-R-A1-es-g-03-0-UI.3' similar to gi|458845|emb|X70497|RNENACA.R.
    norvegicus mRNA for sodium channel, alpha subunit, mRNA sequence.
  ACCESSION AA925917
  VERSION AA925917.1 GI:4237108
  KEYWORDS EST.
  SOURCE Norway rat.
  ORGANISM Rattus norvegicus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
    Rattus.
```

1 (bases 1 to 387)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
On Apr 21, 1998 this sequence version replaced gi:3073053.
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence tag present in the cDNA between the NotI site and the
oligo-dT track served to identify it as a clone from the normalized
adult kidney library. cDNA Library Preparation: M. Fatima Bonaldo,
Ph.D. Clone distribution: clones will be available through Research
Genetics This clone is also available through the I.M.A.G.E.
Consortium at LBNL (info@image.llnl.gov). IMAGE ID=1771734
Seq primer: M13 Forward
PolyA-No.

Location/Qualifiers
1. .387
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-A1-es-g-03-0-UI"
/clone_lib="UI-R-A1"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="vector: pMT30-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-A1
library is a subtracted library derived from the UI-R-A0
library. The UI-R-A0 library consisted of a mixture of
individually tagged normalized libraries constructed from
rat placenta, adult lung, brain, liver, kidney, heart,
spleen, ovary, and muscle. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-A1) was constructed as follows: PCR
amplified cDNA inserts from a pool of approximately 3,840
UI-R-A0 clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-A0
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was
purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-A1 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996)"

BASE COUNT 131 a 102 g 84 t
ORIGIN

Query Match 12.0%; Score 372.6; DB 32; Length 387;
Best Local Similarity 97.7%; Pred. No. 1.3e-82;
Matches 378; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2714 tcatttgatacacaccatccccagtcctctgtggtgtgtctctctctctctctctc 2773
Db 387 TCATTGTATACACATATACCCAGTGCCTCTGTGGTGTGTCTCTCTCTCTCTCTCTC 328
Qy 2774 agcctggagctcttcttactgtccctcagaacacacagctatgatcccttagaaacc 2833
Db 327 AGCTTGGAAAGCTTCTTCACTGTCCCTCAGACACACAGCTTTATGTCATCCCTAGAAACT 268
Qy 2834 tgcacaaaggaccattcttttttgaaactcagttaccctgtattgtctccccagaatt 2893
Db 267 TGCCCAAAAGGCCACCTTTCTTTTGTGAAGTCACTACCTGTATGTTTCCCCCAAGATT 208

```
Qy 2894 gtgcctctccctcccccacagtgactgtattatgttcacattcttttggtc 2953
Db 207 GTGCCCTTCCCTCACCACAGGTACTGTATTATATGTTTCACATCTTTTGTGC 148
Qy 2954 tgcctccctgggtaaacagtcctctgtgtgtcaggatgagattgtctgtttgt 3013
Db 147 TGCTCCCTGGGTAAATGTAGCTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 88
Qy 3014 atccttcgggtctagccagctctccacttgggacggttaggcaggtactcaataaaggc 3073
Db 87 ATCTTCCGGTTAGCCCGACTTCCCACTTGGGACGGGTAGGCAGGTACTCAATAAAGC 28
Qy 3074 ttgtccatcaaaaaaaaaaaaaa 3100
Db 27 TTGTTCCATCAAAAAAAAAAAAAA 1

RESULT 5
AA393950 499 bp mRNA EST 12-AUG-1997
DEFINITION zt78a10.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:728442
5' similar to gb:L29007.cds1 AMILORIDE-SENSITIVE SODIUM CHANNEL
ALPHA-SUBUNIT (HUMAN);, mRNA sequence.
ACCESSION AA393950
VERSION AA393950.1 GI:2046919
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 499)
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,I., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1837 Std Error: 0.00
Seq primer: -28m13 rev2 Et from Amersham
High quality sequence stop: 459.
Location/Qualifiers
1..499
/organism="Homo sapiens"
/db_xref="GDB:592535"
/db_xref="taxon:9606"
/clone="IMAGE:728442"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'.
TGTACCATCTGAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 111 a 138 c 127 g 121 t 2 others
ORIGIN

Query Match 11.2%; Score 348; DB 25; Length 499;
```

```
Best Local Similarity 83.3%; Pred. No. 2e-76;
Matches 408; Conservative 0; Mismatches 80; Indels 2; Gaps 1;
Qy 1427 aggagaacatgatacaagaagtgtgctgtgctacatttaccctaaagcccaaggag 1486
Db 1 AGGAGAGCATGATCAAGAGGTGTGGCTGT--CTACATCTTCTATCCGGCGCCCAAGACG 58
Qy 1487 ttgagtctctgactaccgaaagcagagctctgggctattgtctattataaactgcagg 1546
Db 59 TGGAGTACTGTGACTACAGAAGCACAGTTCCTGGGGGTACTGCTACTTAAAGTCCAGG 118
Qy 1547 ggccttctcttggacagcctgggctgtttctccaaagtgcggaagccttgtagtga 1606
Db 119 TTGACTTCTCTCAGACACCCTGGCTGTTTCACCAAGTCCGGAAGCATCGACGGTGA 178
Qy 1607 tcaactacaaactctctgcggctactcaacggtggccatctgtgaagtcctcaggatgga 1666
Db 179 CCAGCTACCACTCTCTGCTGTGTACTCAGATGGCCCTCGGTGACATCCCAGGAATGGG 238
Qy 1667 tcttcgagatgctctcttgcagacaattacactatttaacacaaagaaacgagttg 1726
Db 239 TCTTCCAGATGCTATCGGACAGACAATTTACACCGTCAACAAGAGAATAATGGAGTGG 298
Qy 1727 caaagctcaacatcttcttcaagagctgaactataaactaatctcggagctctcttct 1786
Db 299 CCAAGTCAACATCTTCTCAAGGAGCTGAACACTACAAAACCAATTTCTGAGTCTCCTCTG 358
Qy 1787 tcaagatggtcagcctctcttccaaacctggcagccagtgagcctgtgttggctcgt 1846
Db 359 TCAGATGTTGTCACCTCTCTCTCCAACTGGCAGCCAGTGGAGCTGTGGTTCGGCTCT 418
Qy 1847 ccgtgctctctgtgtgagatgcggacgctcatcttcgacctcctcgtcgtcacactc 1906
Db 419 CGGTGTTGTCGTGTGGAGATGGCTGAGCTCGCTTTTGACCTCTGCTGTCATCATGTTCC 478
Qy 1907 tcatgctgct 1916
Db 479 TCATGCTGCT 488

RESULT 6
AA249071
LOCUS AW249071 475 bp mRNA EST 07-JAN-2000
DEFINITION 2820891.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820891 5',
mRNA sequence.
ACCESSION AW249071
VERSION AW249071.1 GI:6592064
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 475)
NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
On Jul 7, 1999 this sequence version replaced gi:5866355.
Other_ESTs: 2820891.3prime
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: DCTD/DTF cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
trimming: cross_match from University of Washington Genome Center
PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu
Plate: LLCM5 row: G column: 4
```


Qy 1738 atctcttcaggagctgaactataaaactaatcggagtctc 1780
 Db 43 ATCTTTTCAAAGCTGAACATAATAAACTAAATTCGGAGTCTC 1

RESULT 8

AW354686 383 bp mRNA EST 02-FEB-2000
 LOCUS 36764 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION AW354686
 ACCESSION AW354686
 VERSION AW354686.1 GI:6853676
 KEYWORDS EST.
 SOURCE Bos taurus.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 383)
 Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
 Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and
 Keefe,J.W.

TITLE Design and use of four pooled tissue normalized cDNA libraries for
 EST discovery in cattle

JOURNAL

COMMENT Unpublished (2000)
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACACGCTATGACCAT

BACKWARD: GTTTCCCGCAGTCACGACG

Plate: 21 row: K column: 24

Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

1..383
 Location/Qualifiers
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 2BOV"
 /tissue_type="pooled"
 /lab_host="DH10p"
 /note="Vector: pCMV SPORt6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."
 BASE COUNT 86 a 98 c 114 g 85 t
 ORIGIN

Query Match 9.1%; Score 282.4; DB 71; Length 383;
 Best Local Similarity 83.9%; Pred. No. 4.6e-60;
 Matches 319; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 1218 ggtgcatggtcaggatgagctgctttatgatgatgggttcacacttgagcctgg 1277

Db 4 GGTGCATGAACGGGATGAGCCTGCTTTATGGATGATGCTGGCTTTATCTGCGGCTGG 63

Qy 1278 cgtggagacctccatcagtagagaaaggagccctggagacgctcgagagaaattacgg 1337

Db 64 CFTGGAGACCTCCATCAGCTAGTAGAGGAGCTGTGGACAGACTTGGGGTGACTATGG 123

Qy 1338 cgactgactgagaatggtagcgtatgcccggtcaagaacctttacccttcccaagtatac 1397

Db 124 CGACTGCACCAAGAATGGCAGCGAGGTCGCCAGTGGAGAACCTTTACAAATACCAAGTACAC 183

Qy 1398 acagcaggtgtgattcaactcctgctcccgaggagacatgataagaagtgtggtctgac 1457

Db 184 GCAGCAGGTGTGCTCATTCATTCGTGCTTCCAGGAGAGCATGATCAAGGAGTGTGGCTGTGC 243

Qy 1458 ctacatcttacccttaagcccaaggaggttgagttctgtactacgcgaagacagagctc 1517
 Db 244 CTACATCTTCTACCGCGCGCCGACGCGGTAGAGTTCTGTGACTACAGAGACATAAATTC 303
 Qy 1518 ctggggctattgctattataaaactgcaggcgccctctcctcttgagacagcctgggctgttt 1577
 Db 304 CTGGGGCTACTGCTACTATAAGCTCCAGGATGCCTCTCTCCTCAGACCGCCTGGGCTGTTT 363
 Qy 1578 ctccaagtgtcggagcctt 1597
 Db 364 CACCAAGTGCCTGGAAGCCAT 383

RESULT 9

AI017422/c 382 bp mRNA EST 27-AUG-1998
 LOCUS ou96el2.x1 NCI-CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1635694 3'
 DEFINITION similar to gb:L29007_cds1 AMILORIDE-SENSITIVE SODIUM CHANNEL
 ALPHA-SUBUNIT (HUMAN);, mRNA sequence.

ACCESSION AI017422
 VERSION AI017422.1 GI:3231758
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 382)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>;

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 479 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES

source

1..382
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1635694"
 /clone_lib="NCI-CGAP_Kid3"
 /lab_host="DH10p"
 /note="Organ: kidney; Vector: p7T73D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer,
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified p7T73 vector. mRNA
 source: 2 pooled kidneys. Library went through one round
 of normalization. Library constructed by Bento Soares and
 M. Fatima Bonaldo."
 BASE COUNT 87 a 93 c 96 g 106 t
 ORIGIN

Query Match 8.1%; Score 254; DB 34; Length 382;
 Best Local Similarity 88.7%; Pred. No. 5.8e-53;
 Matches 275; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1038 gaattattccaagttccaccaccctcgtacgggaactgtctacacttttcaatgacaagaa 1097

Db 371 GAATTACTCTCACTCCACCACCGCGATGATGGAACCTGCTATACCTTTCAATGACAAGAA 312

Qy 1098 caactccaatctctggtgtctcctcctgctgagtcgaactggtttgtccctgacact 1157

tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT ORIGIN	73 a	84 c	88 g	70 t
Query Match	7.1%	Score 222.2;	DB 74;	Length 315;
Best Local Similarity	86.6%	Pred. No. 4.9e-45;		
Matches	245;	Conservative 0;	Mismatches 38;	Indels 0;
Gaps	0;			
Qy	1241	cccttatggatgatgggttcaactgagccctgagcctggtgagacacctccatcatgatga	1300	
Db	1	CCCTTATGGATGATGGTGGCTTTAACTTGGCGCCTGGCGTGGAGACCTCCATCGACATGA	60	
Qy	1301	gaaaggaagccctggacagcctcgaggaataattacgcgactgactgagaagaatgtagcg	1360	
Db	61	GGAGGAAACCTGGACAGACTTGGGGCGGATATATGGCGACTGCACCAAGATGGCAGTG	120	
Qy	1361	atgtcccggttcaagaacctttacccttccaaagtatacacagcaggtgtgcaattcactct	1420	
Db	121	ATGTCTCTGTTGAGAACCTTTACCTTCAAAAGTACACACAGCAGGTGTGTATTCACTCT	180	
Qy	1421	gcttccaggagaacatgatacaagaagtgtggtgctgtgctactatcttctaccctaaagccca	1480	
Db	181	GCTTCCAGGAGACATGATCAAGAGAGTGTGGCTGTGCTACATCTTCTATCCGCGCC	240	
Qy	1481	agggagttgagttctgtactaccgaagcagagctctctggg	1523	
Db	241	AGACGTGGAGTACTGTGACTACAGAAAGCACAGTTCTCTCTGGG	283	
RESULT 12				
AW325627				
LOCUS	AW325627	349 bp	mRNA	EST
DEFINITION	17814 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.			27-JAN-2000
ACCESSION	AW325627			
VERSION	AW325627.1	GI:6761548		
KEYWORDS	EST.			
SOURCE	plg.			
ORGANISM	Sus scrofa			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
	1 (bases 1 to 349)			
	Fahrenkrug, S.C., Preking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,			
	Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A. and Keele, J.W.			
	Design and use of two pooled tissue normalized cDNA libraries for			
	EST discovery in swine			
	Unpublished (2000)			
	On May 11, 1999 this sequence version replaced gi:4776412.			
	Contact: Smith TPL			
	USDA, ARS, US Meat Animal Research Center			
	PO Box 166, Clay Center, NE 68933-0166, USA			
	Tel: 402 762 4366			
	Fax: 402 762 4390			
	Email: smith@email.marc.usda.gov			
	Single pass sequencing. Bases called and trimmed with phred			
	v0.980904.e. Vector identified by cross_match with the -minscore 20			
	and -mismatch 12 options.			
	PCR Primers			
	FORWARD: AGGAACAGCTATGACCAT			
	BACKWARD: GTTTCCTCCAGTCACGACG			
	Plate: 8 row: N column: 1			
	Seq primer: ATTTAGTGTGACATATAG.			
FEATURES	Location/Qualifiers			
source	1. .349			
	/organism="Sus scrofa"			
	/db_xref="taxon:9823"			
	/clone_lib="MARC 1P1G"			
	/tissue_type="pooled"			
	/lab_host="Dhi10B"			
	/note="Vector: pCMV SPOR6; Site_1: XbaI; Site_2: XhoI;			
	Library made from pooled tissue from day 11, 13, 15, 20,			
	and 30 embryos."			

BASE COUNT ORIGIN	84 a	118 c	93 g	54 t
Query Match	7.0%	Score 218.4;	DB 71;	Length 349;
Best Local Similarity	80.2%	Pred. No. 4.5e-44;		
Matches	284;	Conservative 0;	Mismatches 61;	Indels 2;
Gaps	0;			
Qy	560	ttaatccttcagatagctgaaataaaagagagctggaagctgagccgcatcacgg	619	
Db	2	TCATFCCCTACAGGTGCAAGGAATTAAGAGGAGCTGGAAGAGCTGGACCGTATCACAG	61	
Qy	620	agcagagcgtcttttgactgtacaaatacaactcttctacactgcgcaggtggtggccc	679	
Db	62	AGCAGACACTCTTCGACCTGTACAAGTACAACCTCTCCAACACTCTCTGCGCCACGCC	121	
Qy	680	gacgccgagctcccgagacctctcgtgtgttttccccacccccctgagcgctgcgca	739	
Db	122	GCCTCCG-----CCGAGACCTGGGGAGCCCTCTGCCACCCCTCTGCAGGCGCTGACGG	175	
Qy	740	ctcacctcccgccctactccggccgacgcgcagcgccagcggtcttc---caggtacgcg	796	
Db	176	TCCGCGCTCCGCCCTCCACTCGGCCCGCAGAGTTCCGAGCGCCACTTCCAGACGCGTGG	235	
Qy	797	acaacaatccccaaagtgcacgggaagactggaagatcgcttccaactgtgcaaccaga	856	
Db	236	ACAACAATCCCGAGTGAAACAGGAGGACTGGAAGATCGGCTTCCAACCTGTGCACACAGA	295	
Qy	857	acaatcagactgtttctaccagacatactctctggtggtgagtcagtgaggg	910	
Db	296	ACAATCAGACTGCTTCTTACCAGACGTACTCATCCGGGTGGATGCAGTGAGGG	349	
RESULT 13				
AA710038				
LOCUS	AA710038	548 bp	mRNA	EST
DEFINITION	vt38g03.r1 Barstead mouse proximal colon MPLRB6 Mus musculus cDNA			24-DEC-1997
ACCESSION	AA710038			
VERSION	AA710038.1	GI:2719956		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1 (bases 1 to 548)			
	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,			
	Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,			
	Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,			
	Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and			
	Waterston, R.			
	The WashU-HMNI Mouse EST Project			
	Unpublished (1996)			
	On Jan 19, 1998 this sequence version replaced gi:2150809.			
	Contact: Marra M/Mouse EST Project			
	WashU-HMNI Mouse EST Project			
	Washington University School of MedicineP			
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			
	Tel: 314 286 1800			
	Fax: 314 286 1810			
	Email: mouseest@wustl.edu			
	This clone is available royalty-free through LLNL; contact the			
	IMAGE Consortium (info@image.llnl.gov) for further information.			
	MGI:631308			
	Seq primer: -28ml3 rev2 ET from Amersham			
	High quality sequence stop: 517.			
FEATURES	Location/Qualifiers			
source	1. .548			
	/organism="Mus musculus"			
	/strain="FVB/N"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:1165396"			
	/clone_lib="Barstead mouse proximal colon MPLRB6"			

/dev_stage="7 day juvenile"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; site_1: EcoRI; site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACGAATCTGAAGTGGGAGGCGCGCCCTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AATTCGGATCCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT7T3 vector.
Library constructed by Bob Barstead."
BASE COUNT 126 a 170 c 127 g 123 t 2 others
ORIGIN

Query Match 7.08; Score 217.4; DB 30; Length 548;
Best Local Similarity 74.2; Pred. No. 9.1e-44;
Matches 386; Conservative 0; Mismatches 83; Indels 52; Gaps 7;
Qy 2175 ctgagagaggagaaggtctctcaccagccctgagctccctgtgtaaacatggtgat 2234
Db 13 CTGAGAGGAGAGTGCTCTCTCACCAGGCCAGTGTCTC-----T 54
Qy 2235 ctcaactcagcacatcttcccccagggcccccgcgagccctctgtgtgtcgcctga 2294
Db 55 GTCACTTTCAGCACATCTTCCACAGCTGCCAG-----CTGCTTTTGGTGTGTCCCGG 107
Qy 2295 gggaacaggtgagtaagggccaggaagccgctccagagg-acaggggctaacgatctg 2353
Db 108 AGTAACAGGCTTAAGCAAGGGGCCAGGAAGTTGTCCAGAGGCAATGGGGCTTAATGAGCTG 167
Qy 2354 ctcaagagctgcctgcctctctctgaacactgctctc--cacaagcacagacaagtc 2411
Db 168 CTCAGAGTGCCTTCCCTGCTCTGACACTGCTTTCACACAAAGACACGGGCAAGTCC 227
Qy 2412 cctctcctcctggatcagccagccagactggagctctgacaggaacttctctggag 2471
Db 228 CCTTTACCTTGGATCAGCCAAAGCCAGACTTGGAGCTTTGACAAGGAACGTTCCCGGGA 287
Qy 2472 aggaccaaggaa-----caaatgaagacgcatcgcactcctac 2510
Db 288 ACGACCAACCAACCGGAACACATATAACAGGCACAGAGAGTGGCCACAGCATTCOCA 347
Qy 2511 cccgtgaccagagagctgctgcctcactgcctcctcagggacactgatctctcctc 2570
Db 348 CCCACGACACAGACTGCGCTGG-CTCACTGCTTTCAAGACACAGATGCTGCTACCC 406
Qy 2571 tccctgaacttgggtggggaacccaccccaaaagcccttggtagttctttggcaattc 2630
Db 407 CTCTTGAACCTGGGTGGGGA--CCACAANNAAGCCCTTGTAGTCTTTTGGCAATTC 464
Qy 2631 cctcctcctcactcactcaggtgggactagagtaagacagac 2673
Db 465 TCCCTTCCCTCACTCCTCTCAGGGTGGGGGCTAGAGTAAGTCTGAC 507

RESULT 14
AA035472 358 bp mRNA EST 10-MAY-1997
LOCUS ZK27a05.r1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
DEFINITION IMAGE:471728 5', similar to gb:L29007_cds1 AMILORIDE-SENSITIVE
SODIUM CHANNEL ALPHA-SUBUNIT (HUMAN);, mRNA sequence.
ACCESSION AA035472
VERSION AA035472.1 GI:1507243
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 358)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,

Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.,
and Marr, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 781 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 214.
Location/Qualifiers
1. 358
/organism="Homo sapiens"
/db_xref="GDB:3757730"
/db_xref="taxon:9606"
/clone="IMAGE:471728"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
AATCGAAGAAATTCGCGCGCCCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."

BASE COUNT 85 a 94 c 81 g 81 t 17 others
ORIGIN

Query Match 6.38; Score 197; DB 20; Length 358;
Best Local Similarity 74.9; Pred. No. 1e-38;
Matches 250; Conservative 0; Mismatches 83; Indels 1; Gaps 1;
Qy 1443 gaagtgtgtgtgcctacatctctaccctaaagcccaagggagtggtgtgacta 1502
Db 6 GCACAGGGGCTGTGCTANCACTTCTATCCGGCGCCCAAGACGTGGAGTACTGTGACTA 65
Qy 1503 ccgaaagcagagctcctgggctattgctattataaactcagggcgctctccttga 1562
Db 66 CAGAAAGCACAGTTCCTGGGGGTACTGCTACTATPAAAGCTCCAGGTGACTTCTCTCAGA 125
Qy 1563 cagcctgggctgttctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1622
Db 126 CCACCTGGCTGTCTTCCAAAGTCCGGAAGCCATGCAGCGTGCACAGTACAGCTCCTC 185
Qy 1623 tgcggctcactcaggggtggccatctgtgaagtcacagattgagattcctcagatctgc 1682
Db 186 TGCTGGTTTACACAGATGGCCCTCGGTGACATCCAGGAATGGGCTTCCAGATGCTATC 245
Qy 1683 cttgcagacaattacactattatacaacaagaacagcaggtgtcgaagctcaacatct- 1741
Db 246 CGCAGAAACAATTAACCGTNAACAACAAGAGNAATNNAAGTGGCCNANTNNAATNTN 305
Qy 1742 tcttcagagagctgaactataaactaatctcga 1775
Db 306 TTTTAAAGGGCTNNACTTTTAAACCCATTNGA 339
RESULT 15
AW105326/c
LOCUS AW105326
DEFINITION xtd60f06.x1 NC1_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2598179 3'

similar to gb:L29007_cds1 AMILORIDE-SENSITIVE SODIUM CHANNEL
ALPHA-SUBUNIT (HUMAN):contains element KER repetitive element ;,
mRNA sequence.
ACCESSION AW105326
VERSION AW105326.1 GI:6076061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 425)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Apr 7, 1998 this sequence version replaced gi:3035002.
Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found
Seq primer: -400P from Gibco
High quality sequence stop: 408.
FEATURES
source
1..425
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2598179"
/clone_lib="NCI-CGAP_Ov23"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.35 kb. Tumor types include: mixed
Mullerian tumor, papillary serous, clear cell, spindle
cell. All are primary tumors, metastasis positive. Life
Technologies catalog #: 11534-013"
BASE COUNT 104 a 118 c 120 g 83 t
ORIGIN

Query Match 5.4%; Score 168.6; DB 63; Length 425;
Best Local Similarity 88.4%; Pred. No. 1.3e-31;
Matches 183; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1141 gggttgctccctgacactgcgcagagcagagcagaaatgacttcacccctgctccacagt 1200
Db 247 GGTCTGCTCCCTGATGCTGCGCGCAGAGCAGAGCAATGACTTCATTCCTCCCTGCTCCACAGTG 188
Qy 1201 acggggccagggtgatgttcacatgagcagatgagcctgctttatgagtggtg 1260
Db 187 ATGGGGCCCCGGGTAATGGTACGGCGCAGGATGAACCTGCTTTATGGATGATGGTGGC 128
Qy 1261 ttcaacttgagcctggcgtggagacccctccatcagatgagaaaggagccctggacagc 1320
Db 127 TTTAACTTGGCGCTGGCGTGGAGACCTCCATCAGCATGAGGAGGAACCCCTGGACAGA 68
Qy 1321 ctcgaggaaattacggcgactgtact 1347
Db 67 CTTGGGGCGATTATGGCGACTGCAC 41

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:56:52 ; Search time 2749.04 Seconds
(without alignments)
2023.578 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcgcaccacgcgtcg.....aaaaaaaaaaaaaaaa 3117

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 972840 seqs, 892348106 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_om.*
- 4: gb_ov.*
- 5: gb_pat.*
- 6: gb_ph.*
- 7: gb_pil.*
- 8: gb_p12.*
- 9: gb_prl.*
- 10: gb_pr2.*
- 11: gb_pr3.*
- 12: gb_ro.*
- 13: gb_sts.*
- 14: gb_sy.*
- 15: gb_un.*
- 16: em_fun.*
- 17: em_hum1.*
- 18: em_hum2.*
- 19: em_in.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_sy.*
- 29: em_un.*
- 30: em_v1.*
- 31: gb_htg1.*
- 32: gb_htg2.*
- 33: gb_in1.*
- 34: gb_in2.*
- 35: em_ba1.*
- 36: em_ba2.*
- 37: em_hum3.*
- 38: em_hum4.*
- 39: gb_pr4.*
- 40: gb_htg3.*
- 41: gb_htg4.*
- 42: gb_htg5.*
- 43: gb_htg6.*

- 44: gb_htg7.*
- 45: em_htg1.*
- 46: em_htg2.*
- 47: em_htg3.*
- 48: em_hum5.*
- 49: gb_pl3.*
- 50: gb_pr5.*
- 51: gb_htg8.*
- 52: gb_htg9.*
- 53: gb_htg10.*
- 54: gb_htg11.*
- 55: gb_htg12.*
- 56: gb_htg13.*
- 57: gb_htg14.*
- 58: gb_in3.*
- 59: gb_htg15.*
- 60: gb_htg16.*
- 61: gb_htg17.*
- 62: em_htg4.*
- 63: em_htg5.*
- 64: em_htg6.*
- 65: em_htg7.*
- 66: em_hum6.*
- 67: gb_htg18.*
- 68: gb_htg19.*
- 69: gb_htg20.*
- 70: gb_htg21.*
- 71: gb_htg22.*
- 72: gb_htg23.*
- 73: gb_htg24.*
- 74: gb_htg25.*
- 75: gb_htg26.*
- 76: gb_htg27.*
- 77: gb_htg28.*
- 78: gb_htg29.*
- 79: gb_htg30.*
- 80: gb_htg31.*
- 81: gb_v11.*
- 82: gb_v12.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3117	100.0	3117	12	RNENAC	X70497 R. norvegic
2	2796	89.7	3081	12	RNASNAC	X70521 R. norvegic
3	2098	67.3	2200	12	RNU54700	U54700 Rattus norv
4	2047	65.7	2200	12	RNU54699	U54699 Rattus norv
5	283	9.1	749	12	AF082073	AF082073 Rattus no
6	283	9.1	2223	12	AF081783	AF081783 Rattus no
7	141	4.5	252	12	RNCDK3UTR	X89818 R. norvegicu
8	104	3.3	3000	12	AF112185	AF112185 Mus muscu
9	98	3.1	98	5	I78503	I78503 Sequence 14
10	71	2.3	852	12	AF002665	AF002665 Rattus no
11	59	1.9	2695	12	CPQ249296	AJ249296 Cavia por
12	45	1.4	45	5	I78500	I78500 Sequence 11
13	45	1.4	45	5	I78501	I78501 Sequence 12
14	41	1.3	2915	3	OCU132108	AJ132108 Oryctolag
15	40	1.3	636	8	PAU95179	U95179 Prunus arme
16	40	1.3	1620	4	GGVITRO	Y11030 G.gallus mr
17	40	1.3	2171	9	AB002533	AB002533 Homo sapi
18	40	1.3	61001	74	AC034204	AC034204 Homo sapi
19	40	1.3	80861	73	AC036192	AC036192 Homo sapi
20	40	1.3	95236	67	AC026450	AC026450 Homo sapi
21	40	1.3	101259	11	HSDJ900E8	AL109623 Human DNA
22	40	1.3	103804	73	AC020928	AC020928 Homo sapi
23	40	1.3	122223	39	AC007880	AC007880 Homo sapi
24	40	1.3	128758	72	AC008932	AC008932 Homo sapi

c	25	40	1.3	141762	10	HS198C21	AL034400 Human DNA	98	1.2	1411	8	ATH011639	AJ011639 Arabidops
c	26	40	1.3	166947	73	AC034243	AC034243 Homo sapi	99	1.2	1525	34	AF212926	AF212926 Pseudale
	27	40	1.3	167627	78	AC015968	AC015968 Homo sapi	100	1.2	1597	33	DDPDEI	X16057 Dictyosteli
	28	40	1.3	170125	41	AC007465	AC007465 Homo sapi	101	1.2	1653	4	AF205877	AF205877 Gallus ga
	29	40	1.3	170273	32	ALI139412	ALI139412 Homo sapi	102	1.2	1686	7	PPPIRNA	X67051 P. patens mr
	30	40	1.3	170948	72	AC012312	AC012312 Homo sapi	103	1.2	1848	8	AF012862	AF012862 Petroselli
	31	40	1.3	172749	39	AC009731	AC009731 Homo sapi	104	1.2	1987	33	DDIRAP1A	L09750 Dictyosteli
c	32	40	1.3	193632	43	AC022037	AC022037 Homo sapi	105	1.2	2050	7	SOAHRI	X57073 S. oliveracea
	33	40	1.3	193731	11	HS80N2	AL031123 Human DNA	106	1.2	2158	5	I17544	I17544 Sequence 1
	34	39	1.3	583	3	BT01B8	X63219 B. taurus CI	107	1.2	2182	12	MMOX1G	Z15103 M. musculus
	35	39	1.3	606	34	AF206698	AF206698 Epiblema	108	1.2	2259	9	AK000421	AK000421 Homo sapi
	36	39	1.3	1845	34	AF123275	AF123275 Dictyoste	109	1.2	2627	25	ATNEMESP	Y11994 Arabidops
	37	39	1.3	2334	9	HUMSODIUM	L29007 Human Kidne	110	1.2	2680	9	AK000323	AK000323 Homo sapi
	38	39	1.3	2562	39	HSASNA	AF060913 Homo sapi	111	1.2	2712	12	AF107352	AF107352 Mus muscu
	39	39	1.3	3151	9	HSASNA	X76180 H. sapiens m	112	1.2	3093	34	DDU77065	U77065 Dictyosteli
	40	39	1.3	4028	5	A39975	A39975 Sequence 8	113	1.2	3175	49	AF029256	AF029256 Kosteletz
	41	39	1.3	4032	5	I23502	I23502 Sequence 8	114	1.2	4126	12	AF102578	AF102578 Mus muscu
	42	39	1.3	6313	10	HS292981	Z292981 H. sapiens S	115	1.2	4254	11	HSM802077	ALI37682 Homo sapi
	43	39	1.3	100368	10	HS514K20	AL008731 Human DNA	116	1.2	4788	33	CBU60649	U60649 Caenorhabdi
	44	39	1.3	140026	41	AC005840	AC005840 Homo sapi	117	1.2	4898	34	DMU07629	U07629 Drosophila
	45	39	1.3	157227	40	ALI158203	ALI158203 Homo sapi	118	1.2	5115	4	AF219232	AF219232 Gallus ga
	46	39	1.3	172081	10	CNS01DVS	ALI136018 Human chr	119	1.2	6691	33	DGSDT1	AJ000992 Dictyoste
	47	39	1.3	188488	39	AC006057	AC006057 Homo sapi	120	1.2	9455	34	AE001378	AE001378 Plasmodiu
	48	39	1.3	194763	56	AC008453	AC008453 Homo sapi	121	1.2	12029	34	AE001408	AE001408 Plasmodiu
	49	38	1.2	75	5	I78504	I78504 Sequence 15	c	1.2	26016	8	SPBC25B2	AL031853 S. pombe c
	50	38	1.2	75	12	S785B2	S785B2 alpha ENaCa	c	1.2	30871	7	SC8021X	Z49704 S. cerevisia
	51	38	1.2	344	12	MMU52006	U52006 Mus musculu	c	1.2	33613	12	AF060868	AF060868 Mus muscu
	52	38	1.2	409	33	DDIAC762	K02959 Slime mold	c	1.2	35005	10	HS30266	AL031703 Human DNA
	53	38	1.2	729	39	AF187015	AF187015 Homo sapi	c	1.2	41906	11	HSAC000124	AC000124 Human Cos
	54	38	1.2	960	5	E02152	E02152 DNA encodin	c	1.2	41947	50	HSAC000358	AC000358 Human Cos
	55	38	1.2	1576	11	MM801341	AL133070 Homo sapi	c	1.2	44560	39	AC004536	AC004536 Homo sapi
	56	38	1.2	2345	12	MMU807620	U87620 Mus musculu	c	1.2	70875	10	HSDJ875H3	AL049733 Human DNA
	57	38	1.2	2851	11	HSM801123	ALI17590 Homo sapi	c	1.2	77835	31	PFMAL13P2_3	Continuation (4 of
	58	38	1.2	3036	11	HSM801281	ALI22118 Homo sapi	c	1.2	77943	39	HSAC002467	AC002467 Human BAC
	59	38	1.2	3152	11	HSM800952	ALI17443 Homo sapi	c	1.2	98845	78	AC004906	AC004906 Homo sapi
	60	38	1.2	3385	12	AF071230	AF071230 Cavia cob	c	1.2	107783	67	AC026445	AC026445 Homo sapi
	61	38	1.2	3586	3	BTU14944	U14944 Bos taurus	c	1.2	108313	32	HSG118G19_3	Continuation (4 of
	62	38	1.2	5078	4	TNU75604	U75604 Tilapia nil	c	1.2	110535	39	AC005091	AC005091 Homo sapi
	63	38	1.2	38705	50	AC000031	AC000031 Homo sapi	c	1.2	120169	39	AC005060	AC005060 Homo sapi
	64	38	1.2	78578	31	PFMAL13P9	AL096783 Plasmodiu	c	1.2	122898	55	AC018366	AC018366 Mus muscu
	65	38	1.2	81562	19	AC004359	AC004359 Drosophil	c	1.2	126088	79	AC044799	AC044799 Homo sapi
	66	38	1.2	137139	10	HSBB455A7	ALI21580 Human DNA	c	1.2	139607	32	HSJ196B13	ALI21968 Homo sapi
	67	38	1.2	140784	54	AC008438	AC008438 Homo sapi	c	1.2	140212	67	AC016790	AC016790 Homo sapi
	68	38	1.2	153009	41	AC009801	AC009801 Homo sapi	c	1.2	148923	68	AC007991	AC007991 Homo sapi
	69	38	1.2	158118	53	AC016537	AC016537 Homo sapi	c	1.2	155236	77	AC009701	AC009701 Homo sapi
	70	38	1.2	158118	53	AC016537	AC016537 Homo sapi	c	1.2	171198	53	AC021036	AC021036 Homo sapi
c	71	38	1.2	172437	59	AC022270	AC022270 Homo sapi	c	1.2	173508	78	AC004898	AC004898 Homo sapi
c	72	38	1.2	174817	43	AC011399	AC011399 Homo sapi	c	1.2	173826	31	AP001160	AP001160 Homo sapi
	73	38	1.2	181693	67	AC026322	AC026322 Homo sapi	c	1.2	174148	73	AC009057	AC009057 Homo sapi
	74	38	1.2	264851	34	AE003644	AE003644 Drosophil	c	1.2	184234	40	ALI162853	ALI162853 Homo sapi
c	75	38	1.2	30962	34	DROSADH04	AE003410 Drosophil	c	1.2	184292	59	AC016763	AC016763 Homo sapi
c	76	37	1.2	320	13	CNS01NFS	ALI52169 Anopheles	c	1.2	180418	44	AC017095	AC017095 Homo sapi
	77	37	1.2	394	14	SYNHET505	J02540 heteropolym	c	1.2	180581	60	AC018609	AC018609 Mus muscu
	78	37	1.2	471	33	AF026265	AF026265 Lucilia c	c	1.2	180798	71	AC011351	AC011351 Homo sapi
	79	37	1.2	495	33	S81785	S81785 Myr p II-ma	c	1.2	183558	32	ALI135792	ALI135792 Homo sapi
c	80	37	1.2	530	13	CNS01TLU	ALI45891 Anopheles	c	1.2	184234	40	ALI162853	ALI162853 Homo sapi
	81	37	1.2	631	8	AF044204	AF044204 Gossypium	c	1.2	186908	31	AP001812	AP001812 Homo sapi
	82	37	1.2	650	11	HSM800760	AL080231 Homo sapi	c	1.2	190877	31	AP001269	AP001269 Homo sapi
	83	37	1.2	727	8	ATU94495	U94495 Arabidops	c	1.2	192581	31	PFMAL13P1	AL049180 Plasmodiu
	84	37	1.2	739	7	ATCTGTGPX2	AJ000470 Arabidops	c	1.2	194932	32	CNS01DXH	ALI139316 Homo sapi
	85	37	1.2	739	39	AF067168	AF067168 Homo sapi	c	1.2	195102	60	AC005506	AC005506 Plasmodiu
	86	37	1.2	858	8	OSU74296	U74296 Oryza sativ	c	1.2	204663	56	AC022105	AC022105 Homo sapi
	87	37	1.2	872	33	DMRNAPB	X75499 D. melanog	c	1.2	204917	39	AC008040	AC008040 Homo sapi
	88	37	1.2	970	5	E00424	E00424 Pig elastas	c	1.2	205804	71	AC027654	AC027654 Mus muscu
	89	37	1.2	996	7	ENASPD1R	Z80229 Emericella	c	1.2	209916	74	AC010422	AC010422 Homo sapi
	90	37	1.2	1001	5	AR016469	AR016469 Sequence	c	1.2	210200	53	AC019238	AC019238 Homo sapi
	91	37	1.2	1012	11	PAU82433	U82433 Prunus arne	c	1.2	214718	79	AC046147	AC046147 Mus muscu
	92	37	1.2	1062	11	HSM802518	ALI62002 Homo sapi	c	1.2	216524	51	AC007305	AC007305 Mus muscu
	93	37	1.2	1107	7	ST27G1433	Y11686 Solanum tub	c	1.2	217304	67	AC010244	AC010244 Homo sapi
	94	37	1.2	1220	7	SOPXR1	Y10462 S. oliveracea	c	1.2	220995	69	AC015995	AC015995 Homo sapi
	95	37	1.2	1223	11	HSU093205	U93205 Homo sapien	c	1.2	225490	32	ALI158052	ALI158052 Homo sapi
	96	37	1.2	1341	33	CCLDOPACB	Y11906 C. capitata	c	1.2	230129	41	HSAC76P10	ALI132867 Homo sapi
	97	37	1.2	1381	7	ATPHDIPSN	X83764 A. thaliana	c	1.2	248619	67	AC008875	AC008875 Homo sapi

171	37	1.2	252102	60	AC005140	244	36	1.2	907	8	ATU95035	U95035 Arabidopsis
172	37	1.2	253307	33	PFMAL3P7	245	36	1.2	912	12	RNY17327	Y17327 Rattus norv
173	37	1.2	256172	41	AC005139	246	36	1.2	918	7	ZMNB1B	X66077 Z.mays MNB1
174	37	1.2	282806	60	AC006279	247	36	1.2	925	7	AB010898	AB010898 Daucus ca
175	37	1.2	308930	67	AC016562	248	36	1.2	925	10	S75940	S75940 (Alu repeat
176	36	1.2	62	105	S2152S31	249	36	1.2	946	49	AF010228	AF010228 Lycopersi
177	36	1.2	105	13	HUMUT956B	250	36	1.2	950	82	AF012335	AF012335 Human End
178	36	1.2	1212	9	AB023798	251	36	1.2	952	8	AF107026	AF107026 Triticum
179	36	1.2	299	39	AF162270	252	36	1.2	958	4	S54890	S54890 growth horm
180	36	1.2	313	13	G36148	253	36	1.2	961	9	AK000718	AK000718 Homo sapi
181	36	1.2	333	9	HSALU006	254	36	1.2	972	5	I22272	I22272 Sequence 1
182	36	1.2	341	9	HSC2N4	255	36	1.2	973	7	VRSA	X70671 V.radiata m
183	36	1.2	345	12	RNRPRL38	256	36	1.2	978	4	AF012462	AF012462 Pleuronec
184	36	1.2	346	9	HSALU005	257	36	1.2	979	7	VVCHI	X75963 V.vinifera
185	36	1.2	346	9	HSC4N2	258	36	1.2	982	7	AB024276	AB024276 Citrus un
186	36	1.2	346	11	HSU67810	259	36	1.2	986	39	AF169154	AF169154 Homo sapi
187	36	1.2	361	10	HUMBRCA2	260	36	1.2	993	8	PRU90347	PRU90347 Pinus radi
188	36	1.2	380	5	A39972	261	36	1.2	1016	34	TU087096	TU087096 Trichomonas
189	36	1.2	380	5	I23499	262	36	1.2	1033	12	MUSUNKNI	MUSUNKNI
190	36	1.2	380	11	S56967	263	36	1.2	1033	12	MUSUNKNI	MUSUNKNI
191	36	1.2	380	11	HSU67832	264	36	1.2	1046	5	A21103	A21103 N.tabacum S
192	36	1.2	397	10	S75201	265	36	1.2	1046	5	I44073	I44073 Sequence 4
193	36	1.2	428	34	AF034637	266	36	1.2	1049	8	AF071889	AF071889 Prunus ar
194	36	1.2	433	5	I93570	267	36	1.2	1050	7	AB029635	AB029635 Daucus ca
195	36	1.2	442	3	OCRPTC	268	36	1.2	1056	12	AF038848	AF038848 Mus muscu
196	36	1.2	462	34	AF155364	269	36	1.2	1061	34	AF001981	AF001981 Dictyoste
197	36	1.2	473	12	RNMDG2	270	36	1.2	1062	7	PALHCB12	PALHCB12
198	36	1.2	490	11	HSU67832	271	36	1.2	1065	3	CHAF000167	CHAF000167 Capra hir
199	36	1.2	500	12	MUSLIBA4	272	36	1.2	1071	11	AF052122	AF052122 Homo sapi
200	36	1.2	503	8	S78730	273	36	1.2	1099	12	AF067790	AF067790 Rattus no
201	36	1.2	514	9	AK000450	274	36	1.2	1103	8	AF020793	AF020793 Medicago
202	36	1.2	514	11	AF015158	275	36	1.2	1105	39	HOMOP2SG1	AF0106543 Homo sapi
203	36	1.2	538	13	G15869	276	36	1.2	1110	11	AF014838	AF014838 Homo sapi
204	36	1.2	541	7	SHY16247	277	36	1.2	1113	5	E12888	E12888 cDNA encodi
205	36	1.2	556	49	AF165420	278	36	1.2	1115	10	MACPOMCA	M19658 Macaca neme
206	36	1.2	565	7	TDEHD25	279	36	1.2	1115	11	HSU96074	U96074 Human trans
207	36	1.2	580	11	HSU67832	280	36	1.2	1130	11	HSU96074	AL137484 Homo sapi
208	36	1.2	590	11	HSU67832	281	36	1.2	1136	8	ASU58949	U58949 Allium sati
209	36	1.2	604	33	HAY12285	282	36	1.2	1138	11	AF026807	AF026807 Homo sapi
210	36	1.2	639	11	HSU67832	283	36	1.2	1149	49	AF159804	AF159804 Vigna unq
211	36	1.2	647	7	D88314	284	36	1.2	1160	9	AK000844	AK000844 Homo sapi
212	36	1.2	667	49	AF339930	285	36	1.2	1162	8	AF117707	AF117707 Lycopersi
213	36	1.2	673	12	RNCEH	286	36	1.2	1173	8	PHRNNAM	X92205 P.hybrida m
214	36	1.2	684	5	AR037084	287	36	1.2	1192	39	AF068235	AF068235 Homo sapi
215	36	1.2	685	7	AB027154	288	36	1.2	1199	39	AF161517	AF161517 Homo sapi
216	36	1.2	701	7	PAB131103	289	36	1.2	1201	49	AF063805	AF063805 Phasodact
217	36	1.2	709	49	AF118113	290	36	1.2	1209	34	AF100287	AF100287 Trypanoso
218	36	1.2	720	34	AF082530	291	36	1.2	1225	7	ZMNBIA	X66076 Z.mays MNB1
219	36	1.2	722	8	AF030252	292	36	1.2	1231	9	AK000476	AK000476 Homo sapi
220	36	1.2	723	7	AB015760	293	36	1.2	1239	12	AF000577	AF000577 Rattus no
221	36	1.2	742	5	I77092	294	36	1.2	1246	7	ATPRXR4GE	X98316 A.thaliana
222	36	1.2	742	7	CACORA16	295	36	1.2	1246	7	PVPFE	X58274 P.vulgaris
223	36	1.2	754	7	RICORFB	296	36	1.2	1247	7	ATFO1	X99096 A.thaliana
224	36	1.2	772	39	MMBGLC2	297	36	1.2	1250	5	A38574	A38574 Sequence 4
225	36	1.2	776	10	HSU46465	298	36	1.2	1255	7	AB028149	AB028149 Medicago
226	36	1.2	778	10	HSU46465	299	36	1.2	1259	5	E03348	E03348 cDNA sequen
227	36	1.2	788	33	AF030184	300	36	1.2	1267	11	HSU800196	AL050109 Homo sapi
228	36	1.2	799	11	HSU67233	301	36	1.2	1273	15	AX001570	AX001570 Sequence
229	36	1.2	799	50	AF151080	302	36	1.2	1276	39	AF028823	AF028823 Homo sapi
230	36	1.2	800	7	PCPR2	303	36	1.2	1278	4	AF100931	AF100931 Oncorhyn
231	36	1.2	806	8	CPA011715	304	36	1.2	1281	11	HSU8011360	AL133084 Homo sapi
232	36	1.2	808	13	G41948	305	36	1.2	1284	5	A20553	A20553 NF-YB. 6/19
233	36	1.2	825	8	AF052871	306	36	1.2	1298	39	AF113689	AF113689 Homo sapi
234	36	1.2	859	11	HSU802545	307	36	1.2	1303	11	HSU90913	U90913 Human clone
235	36	1.2	869	8	LEU84140	308	36	1.2	1316	3	OCU131955	AL131955 Oryctolag
236	36	1.2	884	4	PRDB428	309	36	1.2	1319	39	HUMTRNSAL	AL19437 Homo sapien
237	36	1.2	887	11	HSU801336	310	36	1.2	1321	11	HSU801113	AL117580 Homo sapi
238	36	1.2	893	5	I52156	311	36	1.2	1321	33	TU087094	TU087094 Trichomonas
239	36	1.2	893	5	I52156	312	36	1.2	1329	7	ATH6787	AT006787 Arabidops
240	36	1.2	898	11	HSU801961	313	36	1.2	1332	8	OSHOX1	X96681 Oryza sativ
241	36	1.2	900	4	AB029757	314	36	1.2	1332	11	HSU8000047	AL050389 Homo sapi
242	36	1.2	904	34	AF091604	315	36	1.2	1338	5	AR055519	AR055519 Sequence
243	36	1.2	905	4	GRNAPFCP	316	36	1.2	1352	34	OCU67148	U67148 Osterlagia
									1354	11	HSU801325	AL133054 Homo sapi

317	36	1.2	1362	9	AK000559	AK000559 Homo sapi	390	1.2	1978	11	AF016709	AF016709 Homo sapi
318	36	1.2	1364	7	AB025573	AB025573 Nicotiana	391	1.2	1994	33	AF017442	AF017442 Crassos
319	36	1.2	1366	11	HSB800180	AL050097 Homo sapi	392	1.2	2002	12	MUSBMK	J03023 Murine macr
320	36	1.2	1375	7	EGWWDH	X78900 E.gunnli mR	393	1.2	2014	39	AF122922	AF122922 Homo sapi
321	36	1.2	1380	8	AF038585	AF038585 Zea mays	394	1.2	2018	9	HSXOC11P	AJ000041 Homo sapi
322	36	1.2	1384	12	RATRC3A	L09119 Rattus norv	395	1.2	2019	11	HSB802272	AL137536 Homo sapi
323	36	1.2	1395	5	CA1575	A41575 Sequence 1	396	1.2	2021	7	ATSHAGGY	Y07822 A.thaliana
324	36	1.2	1395	7	CMGTB	X73314 C.maxima mR	397	1.2	2023	8	AF012899	AF012899 Sambucus
325	36	1.2	1402	11	AF003738	AF003738 Homo sapi	398	1.2	2036	7	PSY17796	Y17796 Pisum sativ
326	36	1.2	1430	12	AF058798	AF058798 Mus muscu	399	1.2	2047	8	AF124369	AF124369 Nicotiana
327	36	1.2	1432	10	IKR075868	AL109672 Homo sapi	c 400	1.2	2055	8	AGU79997	U79997 Atriplex ga
328	36	1.2	1435	11	HSB801967	AL137298 Homo sapi	401	1.2	2063	11	HSB801958	U25244 Human hepat
329	36	1.2	1441	5	A65340	A65340 Sequence 63	402	1.2	2064	39	HSU92544	U92544 Human hepat
330	36	1.2	1442	15	AX001574	AX001574 Sequence	403	1.2	2065	11	HSB801264	AL122107 Homo sapi
331	36	1.2	1449	9	AK000392	AK000392 Homo sapi	404	1.2	2082	5	AR059958	AR059958 Sequence
332	36	1.2	1450	9	HS9112	X57348 H.sapiens m	405	1.2	2095	39	AF147790	AF147790 Homo sapi
333	36	1.2	1451	34	AF201906	AF201906 Drosophill	406	1.2	2114	5	I22020	I22020 Sequence 6
334	36	1.2	1487	8	AF169800	AF169800 Lillium hy	407	1.2	2126	9	AK000248	AK000248 Homo sapi
335	36	1.2	1488	9	AK000037	AK000037 Homo sapi	408	1.2	2137	9	AK000251	AK000251 Homo sapi
336	36	1.2	1500	4	XLNAKATP	Y11587 X.laevvis mR	409	1.2	2141	8	AF161885	AF161885 Macadamia
337	36	1.2	1503	11	HSB800520	AL050223 Homo sapi	410	1.2	2145	7	CSNADP	X58832 C.sorokinia
338	36	1.2	1504	11	AR052157	AF052157 Homo sapi	411	1.2	2147	12	MUSP281AM	M25244 Mouse pre-B
339	36	1.2	1505	12	AF215669	AF215669 Mus muscu	412	1.2	2160	12	MMNOD	X70514 M. musculus
340	36	1.2	1534	11	HSB801279	AL122116 Homo sapi	413	1.2	2165	12	MMU96685	U96685 Mus musculu
341	36	1.2	1535	39	AF109683	AL122116 Homo sapi	414	1.2	2186	49	AF139205	AF139205 Abies gra
342	36	1.2	1555	11	HSB802203	AL137478 Homo sapi	415	1.2	2189	11	HSB802178	AL137462 Homo sapi
343	36	1.2	1588	5	E03349	E03349 cDNA sequen	416	1.2	2198	12	AB007812	AB007812 Mus muscu
344	36	1.2	1588	11	HSB800676	AL080159 Homo sapi	417	1.2	2201	12	AF109155	AF109155 Mus muscu
345	36	1.2	1590	34	AF132551	AF132551 Drosophill	418	1.2	2202	39	AF126488	AF126488 Homo sapi
346	36	1.2	1596	33	LFRAC21596	X96512 L. terrestris	c 419	1.2	2220	33	DDCAMP1	X52688 D.discoidleu
347	36	1.2	1602	5	AR0000496	AR0000496 Sequence	420	1.2	2225	7	OSA243828	AJ243828 Orzya sat
c 348	36	1.2	1606	39	AF113015	AF113015 Homo sapi	421	1.2	2245	9	AK000482	AK000482 Homo sapi
c 349	36	1.2	1628	10	HSU08374	U08374 Human cytos	422	1.2	2255	34	AF201905	AF201905 Drosophill
350	36	1.2	1631	8	U79556	U79556 Arabidopsis	423	1.2	2270	11	HSB802322	AL137574 Homo sapi
351	36	1.2	1632	5	E07941	E07941 cDNA encodi	424	1.2	2271	9	HUMASP	J05032 Human aspar
352	36	1.2	1639	5	AR072149	AR072149 Sequence	425	1.2	2297	11	HSB802304	AL137561 Homo sapi
353	36	1.2	1649	8	ATH7450	AJ007450 Arabidops	426	1.2	2334	7	NEUCAM	M73032 Neurospora
354	36	1.2	1670	11	HSB801364	AL133088 Homo sapi	427	1.2	2337	11	HSB801110	AF115758 Homo sapi
355	36	1.2	1691	9	AK000453	AK000453 Homo sapi	428	1.2	2337	39	AF090886	AF090886 Homo sapi
356	36	1.2	1699	11	HSB801509	AL133640 Homo sapi	429	1.2	2333	5	A92394	A92394 Sequence 19
357	36	1.2	1699	11	HSU39656	U39656 Human MAP k	430	1.2	2333	7	ATSGT	283833 Arabidopsi
358	36	1.2	1700	1	BRE237608	AJ237608 Borrelia	431	1.2	2368	34	AF190745	AF190745 Drosophill
359	36	1.2	1704	39	HSU81232	U81232 Human tumor	432	1.2	2381	5	I30339	I30339 Sequence 9
360	36	1.2	1723	49	AF136004	AF136004 Triticum	433	1.2	2381	5	I96063	I96063 Sequence 9
361	36	1.2	1725	8	AF007785	AF007785 Zea mays	434	1.2	2384	5	I30334	I30334 Sequence 1
362	36	1.2	1748	11	AF054988	AF054988 Homo sapi	435	1.2	2384	5	I96058	I96058 Sequence 1
363	36	1.2	1776	7	RCENOLASE	Z28386 R.communis	436	1.2	2394	11	AF061795	AF061795 Homo sapi
364	36	1.2	1792	8	AF170272	AF170272 Manihot e	c 437	1.2	2409	7	SCYNL142W	Y1418 S.cerevisia
365	36	1.2	1801	12	MMU6584	AJ006584 Mus muscu	438	1.2	2421	10	HSAR1	Y14314 Homo sapien
366	36	1.2	1804	33	ACU89984	U89984 Acanthamoeb	439	1.2	2440	12	RNU39207	U39207 Rattus norv
367	36	1.2	1808	9	HSOMGPA	X51694 H.sapiens O	440	1.2	2445	8	AF019113	AF019113 Orzya sat
368	36	1.2	1817	8	ATJ00163	AJ001363 Arabidops	441	1.2	2472	39	AF151685	AF151685 Homo sapi
369	36	1.2	1838	39	AF206502S2	AF206503 Homo sapi	442	1.2	2475	11	HSB801926	AL137266 Homo sapi
370	36	1.2	1846	39	AF090940	AF090940 Homo sapi	443	1.2	2486	33	CEU23830	U23830 Caenorhabdi
371	36	1.2	1857	11	AF007142	AF007142 Homo sapi	444	1.2	2499	11	HSB801350	AL133016 Homo sapi
372	36	1.2	1872	9	HSCHRX	X65724 H.sapiens D	c 445	1.2	2500	34	DDU89350	U89350 Dictyosteli
373	36	1.2	1882	12	RNCALRET	X53363 Rat mRNA fo	446	1.2	2530	11	HSB801333	AL133062 Homo sapi
c 374	36	1.2	1891	9	HUMENTRALU	L30117 Human heat-	c 447	1.2	2532	33	SLMACTARD	M15272 Sline mold
375	36	1.2	1893	11	AF070573	AF070573 Homo sapi	448	1.2	2563	7	NEUGPROTAA	L11452 Neurospora
376	36	1.2	1899	12	AF093542	AF093542 Mus muscu	449	1.2	2568	33	DDP8A7	X51947 Dictyosteli
377	36	1.2	1915	11	AF038191	AF038191 Homo sapi	450	1.2	2575	11	HSB801461	AL133653 Homo sapi
378	36	1.2	1929	7	LEGALK	X99851 L.erecta mR	451	1.2	2648	11	HSB801470	AL133656 Homo sapi
379	36	1.2	1930	12	MMU249198	AJ249198 Mus muscu	452	1.2	2675	11	HSB801280	AL122117 Homo sapi
380	36	1.2	1934	4	S59519	S59519 stannioalc	453	1.2	2718	12	AF199027	AF199027 Mus muscu
381	36	1.2	1934	9	AK000074	AK000074 Homo sapi	454	1.2	2727	11	HSB801383	AL133103 Homo sapi
382	36	1.2	1942	9	HSB8010841	AJ010841 Homo sapi	455	1.2	2762	11	AF039916	AF039916 Homo sapi
383	36	1.2	1945	33	CEU61236	U61236 Caenorhabdi	456	1.2	2763	11	HSB802278	AL137541 Homo sapi
384	36	1.2	1952	5	I28450	I28450 Sequence 1	457	1.2	2816	11	HSB800549	AL117440 Homo sapi
385	36	1.2	1952	5	I81145	I81145 Sequence 1	458	1.2	2876	11	HSB800660	AL080147 Homo sapi
386	36	1.2	1952	5	I84698	I84698 Sequence 1	459	1.2	2896	4	AF061981	AF061981 Xenopus l
387	36	1.2	1957	12	AF098295	AF098295 Mus muscu	460	1.2	2922	11	HSB802222	AL137495 Homo sapi
c 388	36	1.2	1964	7	SCYNL141W	Y14117 S.cerevisia	461	1.2	2923	7	ATHSP91	Z70314 A.thaliana
389	36	1.2	1968	10	HSU09115	U09115 Human retro	462	1.2	2924	11	HSB800664	AL080150 Homo sapi

463	36	1.2	2924	49	AF188678	AF188678 Pisum sat	c 536	36	1.2	12593	34	AE001418	AE001418 Plasmodiu
464	36	1.2	2939	11	HSM801345	AL133074 Homo sapi	c 537	36	1.2	12597	39	HSCSE1G3	AF053645 Homo sapi
465	36	1.2	2966	5	A271171	A271171 Tyrosine ki	c 538	36	1.2	12661	11	AL133345	AL133345 Human DNA
466	36	1.2	2973	9	HUMSINEE	D14548 Human DNA,	c 539	36	1.2	12728	12	RNU22062	U22062 Rattus norv
467	36	1.2	2990	50	AF138859	X99971 Macaca fasc	c 540	36	1.2	12896	12	AB0101503	U22062 Mus muscu
468	36	1.2	3015	10	MEDAZMRNA	AF138859 Homo sapi	c 541	36	1.2	13121	12	MMSRP20	X91636 M.musculus
469	36	1.2	3024	39	AF146568	AF146568 Homo sapi	c 542	36	1.2	13208	11	AB028893	AB028893 Homo sapi
470	36	1.2	3078	4	AF225906	AF225906 Xenopus l	c 543	36	1.2	13359	11	HSU89387	U89387 Human RNA p
471	36	1.2	3078	11	HSM801318	AL133047 Homo sapi	c 544	36	1.2	13728	11	HSCDCREL1	AF005988 Homo sapi
472	36	1.2	3113	12	AF056324	AF056324 Rattus no	c 545	36	1.2	14339	34	AE001414	AE001414 Plasmodiu
473	36	1.2	3133	11	HSM800556	AL080057 Homo sapi	c 546	36	1.2	14601	10	HSDJ843G2	AL096762 Human DNA
474	36	1.2	3152	39	HSU78525	U78525 Homo sapien	c 547	36	1.2	14694	10	HUMDZA2G	D14034 Homo sapien
475	36	1.2	3169	11	HSW800953	AL117444 Homo sapi	c 548	36	1.2	15471	8	SPAC1B1	Z98532 S.pombe chr
476	36	1.2	3171	4	AF115497	AF115497 Xenopus l	c 549	36	1.2	15936	10	AB025285	AB025285 Homo sapi
477	36	1.2	3184	10	HS385E7	AL031720 Human DNA	c 550	36	1.2	16471	10	HS889J22A	AL049850 Human DNA
478	36	1.2	3209	34	AF181629	AF181629 Drosophil	c 551	36	1.2	17317	9	AP000250	AP000250 Homo sapi
479	36	1.2	3251	11	HSM800661	AL080148 Homo sapi	c 552	36	1.2	22738	9	AB014078	AB014078 Homo sapi
480	36	1.2	3256	50	AF208850	AF208850 Homo sapi	c 553	36	1.2	23996	71	AC027289	AC027289 Homo sapi
481	36	1.2	3262	33	DME011578	AJ011578 Drosophil	c 554	36	1.2	24159	50	AC005677	AC005677 Homo sapi
482	36	1.2	3336	8	AF118843	AF118843 Lycopersi	c 555	36	1.2	26106	39	AC006005	AC006005 Homo sapi
483	36	1.2	3357	9	HUMIRIT	M64936 Homo sapien	c 556	36	1.2	27170	9	AP000319	AP000319 Homo sapi
484	36	1.2	3363	33	CEMCE1A	X75564 C.elegans m	c 557	36	1.2	27755	9	AP000569	AP000569 Homo sapi
485	36	1.2	3441	11	AF057280	AF057280 Homo sapi	c 558	36	1.2	28247	10	HS362J20	AL049764 Human DNA
486	36	1.2	3451	5	E13998	E13998 Cyanidium c	c 559	36	1.2	28835	10	HS444G9	Z98258 Human DNA s
487	36	1.2	3504	39	AF074606	AF074606 Homo sapi	c 560	36	1.2	29408	41	AC009804	AC009804 Homo sapi
488	36	1.2	3535	12	AF120492	AF120492 Rattus no	c 561	36	1.2	29848	39	U73023	U73023 Homo sapien
489	36	1.2	3551	10	HUM49DC62Z	L77073 Homo sapien	c 562	36	1.2	29952	10	HSV1164A6	Z69838 Human Chrom
490	36	1.2	3581	12	RNSTOP	X93495 R.norvegicu	c 563	36	1.2	30438	9	AP000299	AP000299 Homo sapi
491	36	1.2	3642	11	HSW801377	AL133021 Homo sapi	c 564	36	1.2	33595	32	HS312E8	AL032819 Homo sapi
492	36	1.2	3657	12	MMU51126	U51126 Mus musculu	c 565	36	1.2	33760	10	HSNFG9	Z697119 Human DNA s
493	36	1.2	3676	11	HSW802430	AL157448 Homo sapi	c 566	36	1.2	33905	55	AC010508	AC010508 Homo sapi
494	36	1.2	3700	2	HIU38617	U38617 Haemophilus	c 567	36	1.2	34352	11	U73638	U73638 Human Chrom
495	36	1.2	3738	11	HSW801291	AL133026 Homo sapi	c 568	36	1.2	35173	11	AL160413	AL160413 Human DNA
496	36	1.2	3765	11	HSW800941	AL117432 Homo sapi	c 569	36	1.2	35488	11	AC004589	AC004589 Homo sapi
497	36	1.2	3787	11	HSW801412	AL133565 Homo sapi	c 570	36	1.2	35730	10	HSN74G7	Z697115 Human DNA s
498	36	1.2	3798	9	HSSTREP	X58156 H.sapiens s	c 571	36	1.2	35843	10	HS108F12	Z49235 Human DNA s
499	36	1.2	3820	5	E02221	E02221 cDNA encodi	c 572	36	1.2	36003	51	AC020810	AC020810 Mus muscu
500	36	1.2	3851	5	E01364	E01614 DNA sequenc	c 573	36	1.2	36429	10	HS14B7	Z49258 Human DNA s
501	36	1.2	3851	5	E13364	E13364 Plasmid pyg	c 574	36	1.2	36565	11	HSU73648	Z49258 Human Chrom
502	36	1.2	3851	5	E13540	E13540 Seriola sp.	c 575	36	1.2	36649	11	HSU19D8	Z70689 Human DNA s
503	36	1.2	3970	12	AF027180	AF027180 Cavia por	c 576	36	1.2	36655	10	HS399E4	AL031721 Human DNA
504	36	1.2	4089	10	HUM25DC32	L78753 Homo sapien	c 577	36	1.2	36669	10	HS433G19	AL008735 Human DNA
505	36	1.2	4153	11	HSW801284	AL122121 Homo sapi	c 578	36	1.2	36781	9	AP000304	AP000304 Homo sapi
506	36	1.2	4178	10	HUMCPLA1	D38178 Human gene	c 579	36	1.2	37389	55	AC008974	AC008974 Homo sapi
507	36	1.2	4198	12	MUSHEX	J05277 Mouse hexok	c 580	36	1.2	37658	10	HS366L4	AL023494 Human DNA
508	36	1.2	4215	12	MMU73378	U73378 Mus musculu	c 581	36	1.2	37930	11	AC005564	AC005564 Homo sapi
509	36	1.2	4235	10	HSWEMD	Y10183 H.sapiens m	c 582	36	1.2	38194	39	AC004193	AC004193 Homo sapi
510	36	1.2	4295	9	AK000160	AK000160 Homo sapi	c 583	36	1.2	38357	10	HS462D8	AL022332 Human DNA
511	36	1.2	4417	11	AF098162	AF098162 Homo sapi	c 584	36	1.2	38437	9	AP000361	AP000361 Homo sapi
512	36	1.2	4493	11	HSW801283	AL122120 Homo sapi	c 585	36	1.2	38468	39	U51560	U51560 Homo sapien
513	36	1.2	4579	33	DME269539	AJ269539 Drosophil	c 586	36	1.2	38748	10	HSJ955M13	AL050404 Human DNA
514	36	1.2	4586	5	AR038747	AR038747 Sequence	c 587	36	1.2	38772	10	HSQ19A	Z81310 Human DNA s
515	36	1.2	4586	5	AR059633	AR059633 Sequence	c 588	36	1.2	38968	11	AC005793	AC005793 Homo sapi
516	36	1.2	4740	10	HUMMAC1A	J03525 Human Mac-1	c 589	36	1.2	39000	10	HSN86D4	Z82250 Human DNA s
517	36	1.2	4827	11	AF059611	AF059611 Homo sapi	c 590	36	1.2	39019	10	D88270	D88270 Homo sapien
518	36	1.2	4878	5	I36871	I36871 Sequence 1	c 591	36	1.2	39021	39	AF052041	AF052041 Homo sapi
519	36	1.2	4978	23	E09332	E09332 cDNA encodi	c 592	36	1.2	39144	73	AC008980	AC008980 Homo sapi
520	36	1.2	4978	23	E10112	E10112 Cattle cDNA	c 593	36	1.2	39206	75	AC010649	AC010649 Homo sapi
521	36	1.2	5055	41	AC005843	AC005843 Homo sapi	c 594	36	1.2	39212	10	HSN21F1	Z94162 Human DNA s
522	36	1.2	5233	8	AF020288	AF020288 Arabidops	c 595	36	1.2	39238	39	AC002082	AC002082 Homo sapi
523	36	1.2	5301	11	AF061189	AF061189 Homo sapi	c 596	36	1.2	39239	11	AC005790	AC005790 Homo sapi
524	36	1.2	5307	11	AF040628	AF040628 Homo sapi	c 597	36	1.2	39265	55	AC010505	AC010505 Homo sapi
525	36	1.2	5321	11	HSW802199	AL137476 Homo sapi	c 598	36	1.2	39266	11	AC005791	AC005791 Homo sapi
526	36	1.2	5515	50	AF227198	AF227198 Homo sapi	c 599	36	1.2	39338	11	AC004644	AC004644 Homo sapi
527	36	1.2	5878	39	AF201468	AF201468 Homo sapi	c 600	36	1.2	39372	41	AC004176	AC004176 Homo sapi
528	36	1.2	8200	39	AF126531	AF126531 Homo sapi	c 601	36	1.2	39563	11	HSU73629	U73629 Human chrom
529	36	1.2	8212	12	MMHYLTK	XS8972 M.musculus	c 602	36	1.2	39595	11	AC004509	AC004509 Homo sapi
530	36	1.2	8622	50	AF135026	AF135026 Homo sapi	c 603	36	1.2	39699	11	AC004170	AC004170 Homo sapi
531	36	1.2	8898	55	AC010551	AC010551 Homo sapi	c 604	36	1.2	39834	9	AF001410	AF001410 Homo sapi
532	36	1.2	9404	9	HUMBHSD	M38180 Human 3-bet	c 605	36	1.2	39838	39	AC007766	AC007766 Homo sapi
533	36	1.2	11597	9	AP000693	AP000693 Homo sapi	c 606	36	1.2	39842	9	AP000458	AP000458 Homo sapi
534	36	1.2	11807	9	D86590	D86590 Homo sapien	c 607	36	1.2	39976	11	AC005952	AC005952 Homo sapi
535	36	1.2	11967	10	HUMDNL1L	L40817 Homo sapien	c 608	36	1.2	40024	38	HSEWE5	Y08806 H.sapiens E

c 609	36	1.2	40160	31	AP001214	AP001214 Homo sapi	682	36	1.2	57939	51	AC023608	AC023608 Mus muscu
c 610	36	1.2	40205	11	AF045450	AF045450 Homo sapi	683	36	1.2	58190	10	HS498124	U0751057 Human DNA
c 611	36	1.2	40209	67	AC026805	AC026805 Homo sapi	c 684	36	1.2	59012	11	HSABLGR2	U07562 Human ABL g
c 612	36	1.2	40307	9	AP001430	AP001430 Homo sapi	685	36	1.2	59716	31	AP000637	AF000637 Homo sapi
c 613	36	1.2	40607	11	AC005330	AC005330 Homo sapi	686	36	1.2	59721	51	AC023413	AC023413 Homo sapi
c 614	36	1.2	40616	55	AC016630	AC016630 Homo sapi	687	36	1.2	59850	42	AC016452	AC016452 Homo sapi
c 615	36	1.2	40649	11	AC003111	AC003111 Human DNA	c 688	36	1.2	60129	39	AC006325	AC006325 Homo sapi
c 616	36	1.2	40926	10	HSB10B1	737979 Human DNA s	689	36	1.2	60141	76	AC055858	AC055858 Homo sapi
c 617	36	1.2	41133	11	AC004598	AC004598 Homo sapi	c 690	36	1.2	60179	41	AF149786	AF149786 Homo sapi
c 618	36	1.2	41191	55	AC009004	AC009004 Homo sapi	c 691	36	1.2	60442	9	AP001342	AP001342 Homo sapi
c 619	36	1.2	41495	11	AC004224	AC004224 Homo sapi	692	36	1.2	60801	39	AC006561	AC006561 Homo sapi
c 620	36	1.2	41617	11	AC004213	AC004213 Homo sapi	693	36	1.2	61711	55	AC011530	AL011530 Homo sapi
c 621	36	1.2	41617	11	AC004213	AC004213 Homo sapi	694	36	1.2	62732	10	HS414D7	AL013543 Human DNA
c 622	36	1.2	41696	11	AC005932	AC005932 Homo sapi	c 695	36	1.2	62732	10	HS414D7	AL013543 Human DNA
c 623	36	1.2	41702	11	AC004156	AC004156 Homo sapi	696	36	1.2	62838	78	AC061995	AC061995 Homo sapi
c 624	36	1.2	41726	70	AC027346	AC027346 Homo sapi	c 697	36	1.2	62873	9	AP000238	AP000238 Homo sapi
c 625	36	1.2	41907	39	AF050154	AF050154 Homo sapi	698	36	1.2	62915	75	AC040984	AC040984 Homo sapi
c 626	36	1.2	42179	39	HS076377	U76377 Human olfac	699	36	1.2	63369	41	AC002413	AC002413 Homo sapi
c 627	36	1.2	42345	55	AC010504	AC010504 Homo sapi	700	36	1.2	64063	40	AL353578	AL353578 Homo sapi
c 628	36	1.2	42547	39	AC023172	AC023172 Homo sapi	701	36	1.2	64167	10	HSBA89B2	AL080245 Human DNA
c 629	36	1.2	42558	50	AC037199	AC037199 Homo sapi	c 702	36	1.2	64167	10	HSBA89B2	AL080245 Human DNA
c 630	36	1.2	42601	11	AC004210	AC004210 Homo sapi	703	36	1.2	64767	11	U69730	U69730 Homo sapien
c 631	36	1.2	42649	50	AC024077	AC024077 Homo sapi	c 704	36	1.2	65346	67	AC026597	AC026597 Homo sapi
c 632	36	1.2	42896	10	HS447C4	AF021977 Human DNA	c 705	36	1.2	65730	42	AC014285	AC014285 Drosophil
c 633	36	1.2	42930	50	AF241732	AF241732 Homo sapi	706	36	1.2	66109	9	HSMHCAPG	X66401 H.sapiens g
c 634	36	1.2	43394	50	HSAC000359	AC000359 Human cos	c 707	36	1.2	66739	69	AC008656	AC008656 Homo sapi
c 635	36	1.2	43481	7	SCCHXIV43	Z46843 S.cerevisia	c 708	36	1.2	66741	10	HS460J8	AL031662 Human DNA
c 636	36	1.2	43600	11	AC004209	AC004209 Homo sapi	c 709	36	1.2	68663	73	AC036112	AC036112 Homo sapi
c 637	36	1.2	43689	53	AC022516	AC022516 Homo sapi	710	36	1.2	67053	9	AP001412	AP001412 Homo sapi
c 638	36	1.2	43761	10	HS597B2	AL031346 Human DNA	711	36	1.2	67298	31	AP0000850	AP0000850 Homo sapi
c 639	36	1.2	43761	10	HS597B2	AL031346 Human DNA	c 712	36	1.2	67331	55	AC010286	AC010286 Homo sapi
c 640	36	1.2	43795	11	AC000022	AC000022 Genomic s	713	36	1.2	67939	76	AC055843	AC055843 Homo sapi
c 641	36	1.2	43978	74	AC010617	AC010617 Homo sapi	c 714	36	1.2	67994	39	AC006981	AC006981 Homo sapi
c 642	36	1.2	44210	55	AC016629	AC016629 Homo sapi	c 715	36	1.2	68245	42	AC018626	AC018626 Homo sapi
c 643	36	1.2	44348	11	AC006046	AC006046 Homo sapi	c 716	36	1.2	68581	9	AP000563	AP000563 Homo sapi
c 644	36	1.2	44496	11	AC004760	AC004760 Homo sapi	c 717	36	1.2	68730	79	AC064810	AC064810 Homo sapi
c 645	36	1.2	45333	10	HS10618A	AL049773 Human DNA	718	36	1.2	69015	39	HSU95739	U95739 Human chrom
c 646	36	1.2	44837	55	AC011544	AC011544 Homo sapi	719	36	1.2	69017	10	HS410I8	AB011732 Human DNA
c 647	36	1.2	44890	50	AC005778	AC005778 Homo sapi	c 720	36	1.2	69142	7	AB012239	AB012239 Arabidops
c 648	36	1.2	44890	50	AC005778	AC005778 Homo sapi	c 721	36	1.2	69279	76	AC046187	AC046187 Homo sapi
c 649	36	1.2	44978	50	AF241731	AF241731 Homo sapi	c 722	36	1.2	69517	73	AC036117	AC036117 Homo sapi
c 650	36	1.2	45084	11	AC003005	AC003005 Human DNA	c 723	36	1.2	69529	55	AC012632	AC012632 Homo sapi
c 651	36	1.2	45333	10	HS10618A	AL049773 Human DNA	724	36	1.2	69572	11	AC004769	AC004769 Homo sapi
c 652	36	1.2	45437	41	AC002095	AC002095 Homo sapi	725	36	1.2	69684	39	AC006475	AC006475 Homo sapi
c 653	36	1.2	45550	39	AC006292	AC006292 Homo sapi	c 726	36	1.2	70165	74	AC037463	AC037463 Homo sapi
c 654	36	1.2	46201	39	AC006139	AC006139 Homo sapi	c 727	36	1.2	70488	42	AC015519	AC015519 Homo sapi
c 655	36	1.2	46275	11	AC003107	AC003107 Human DNA	c 728	36	1.2	70739	67	AC026366	AC026366 Homo sapi
c 656	36	1.2	46534	10	HS1057D18	AL035450 Human DNA	c 729	36	1.2	70851	39	AC002540	AC002540 Human BAC
c 657	36	1.2	48092	76	AC053543	AC053543 Homo sapi	c 730	36	1.2	70932	9	AP000269	AP000269 Homo sapi
c 658	36	1.2	48425	51	AC023310	AC023310 Homo sapi	c 731	36	1.2	71534	76	AC046188	AC046188 Homo sapi
c 659	36	1.2	48545	71	AC011515	AC011515 Homo sapi	c 732	36	1.2	71613	73	AC036139	AC036139 Homo sapi
c 660	36	1.2	48566	69	AC025172	AC025172 Homo sapi	733	36	1.2	72289	11	HS091115	HS091115 Human DNA
c 661	36	1.2	48645	11	AC005575	AC005575 Homo sapi	734	36	1.2	72852	31	AP000686	AP000686 Homo sapi
c 662	36	1.2	49261	11	HSBA209A2	AF184110 Homo sapi	735	36	1.2	72882	11	AC004447	AC004447 Homo sapi
c 663	36	1.2	49499	9	AF0184110	AF0184110 Homo sapi	736	36	1.2	73020	11	AC009776	AC009776 Homo sapi
c 664	36	1.2	50188	9	AP000047	AP000047 Homo sapi	c 737	36	1.2	73220	41	AC009776	AC009776 Homo sapi
c 665	36	1.2	50508	39	AC004940	AC004940 Homo sapi	c 738	36	1.2	73231	39	AC005078	AC005078 Homo sapi
c 666	36	1.2	50592	31	AP001236	AP001236 Homo sapi	739	36	1.2	73842	57	AC025657	AC025657 Homo sapi
c 667	36	1.2	50604	39	AC005074	AC005074 Homo sapi	740	36	1.2	74116	42	AC016110	AC016110 Homo sapi
c 668	36	1.2	50768	39	AF095901	AF095901 Homo sapi	c 741	36	1.2	74155	11	AC004656	AC004656 Homo sapi
c 669	36	1.2	50768	39	AF095901	AF095901 Homo sapi	742	36	1.2	74195	72	AC034125	AC034125 Homo sapi
c 670	36	1.2	51859	10	HSJ1193N1	AL121915 Human DNA	c 743	36	1.2	74549	10	HSJ890015	HSJ890015 Human DNA
c 671	36	1.2	53168	10	HS109G5	AL023879 Human DNA	c 744	36	1.2	74729	10	HS703H14	AL031287 Human DNA
c 672	36	1.2	53897	11	AC002037	AC002037 Human Chr	c 745	36	1.2	74817	9	AP001434	AP001434 Homo sapi
c 673	36	1.2	55520	10	HS04488M4	AL078598 Human DNA	746	36	1.2	75270	11	AF023268	AF023268 Homo sapi
c 674	36	1.2	55917	39	AC006002	AC006002 Homo sapi	747	36	1.2	75976	73	AC026633	AC026633 Homo sapi
c 675	36	1.2	56870	39	AC006314	AC006314 Homo sapi	748	36	1.2	76079	39	AC004942	AC004942 Homo sapi
c 676	36	1.2	56837	39	HSMH3A5	U89335 Homo sapien	749	36	1.2	76815	75	AC044789	AC044789 Homo sapi
c 677	36	1.2	56846	46	AC010850	AC010850 Homo sapi	c 750	36	1.2	76822	72	AC032041	AC032041 Homo sapi
c 678	36	1.2	56913	10	HS67C13	280896 Human DNA s	751	36	1.2	76988	39	AC006578	AC006578 Homo sapi
c 679	36	1.2	56997	32	CNS01RG2	AL157690 Homo sapi	752	36	1.2	76990	31	AP001770	AP001770 Homo sapi
c 680	36	1.2	57166	39	AC009320	AC009320 Homo sapi	c 753	36	1.2	77183	10	HS299C21	299290 Human DNA s
c 681	36	1.2	57166	39	AC009320	AC009320 Homo sapi	754	36	1.2	77239	57	AC025680	AC025680 Homo sapi

c 755	1.2	77879	70	AC027403	AC027403 Homo sapi	828	36	1.2	90357	61	AC023880	AC023880 Homo sapi
c 756	1.2	78693	10	HSJD475B7	AL050306 Human DNA	829	36	1.2	90433	40	AL139289	AL139289 Homo sapi
c 757	1.2	79085	39	AC006221	AC006221 Homo sapi	830	36	1.2	90618	57	AC018667	AC018667 Homo sapi
c 758	1.2	79205	75	AC008598	AC008598 Homo sapi	c 831	36	1.2	90744	10	HS3646B12	AL096776 Human DNA
c 759	1.2	79716	11	AC004502	AC004502 Homo sapi	c 832	36	1.2	90815	67	AC002559	AC002559 Homo sapi
c 760	1.2	79868	70	AC027408	AC027408 Homo sapi	c 833	36	1.2	90879	67	AC026356	AC026356 Homo sapi
c 761	1.2	80141	9	HSEWSGAR	Y07848 Homo sapien	c 834	36	1.2	91047	67	AC026430	AC026430 Homo sapi
c 762	1.2	80165	71	AC027745	AC027745 Homo sapi	c 835	36	1.2	91128	31	AP001826	AP001826 Homo sapi
c 763	1.2	80168	39	AC002078	AC002078 Human BAC	c 836	36	1.2	91365	10	HS116F5	Z93244 Human DNA s
c 764	1.2	80209	32	CNS01DVI	AL136294 Homo sapi	c 837	36	1.2	91388	39	AC005994	AC005994 Homo sapi
c 765	1.2	80331	39	AC004994	AC004994 Homo sapi	c 838	36	1.2	91511	11	AL133404	AL133404 Human DNA
c 766	1.2	80312	52	AC007834_3	Continuatiun (4 of	c 839	36	1.2	91634	39	AC005008	AC005008 Homo sapi
c 767	1.2	80518	31	PFMAL13PA	AL109815 Plasmodiu	c 840	36	1.2	91638	41	AC002489	AC002489 Mus muscu
c 768	1.2	80858	39	AC007360	AC007360 Homo sapi	c 841	36	1.2	91638	41	AC002489	AC002489 Mus muscu
c 769	1.2	80869	10	HSJD454N4	AL096794 Human DNA	c 842	36	1.2	91767	39	AC007748	AC007748 Homo sapi
c 770	1.2	81290	51	AC022597	AC022597 Homo sapi	c 843	36	1.2	91886	32	AL133472	AL133472 Homo sapi
c 771	1.2	81340	43	AC021612	AC021612 Homo sapi	c 844	36	1.2	92073	72	AC008968	AC008968 Homo sapi
c 772	1.2	81374	11	HSJ300H18	AL132799 Human DNA	c 845	36	1.2	92357	10	HSJ106901	AL049633 Human DNA
c 773	1.2	81404	52	AC008180_3	Continuatiun (4 of	c 846	36	1.2	92636	39	AC005251	AC005251 Homo sapi
c 774	1.2	81479	70	AC026704	AC026704 Homo sapi	c 847	36	1.2	92651	39	AC005257	AC005257 Human BAC
c 775	1.2	81521	39	AC006255	AC006255 Homo sapi	c 848	36	1.2	92797	11	AF038458	AF038458 Homo sapi
c 776	1.2	81561	9	AB023057	AB023057 Homo sapi	c 849	36	1.2	92817	11	HSJ393D12	AL132776 Human DNA
c 777	1.2	81807	72	AC011433	AC011433 Homo sapi	c 850	36	1.2	92969	11	HSJ687F11	AL117334 Human DNA
c 778	1.2	82116	9	AP000338	AP000338 Homo sapi	c 851	36	1.2	93273	39	AC004084	AC004084 Homo sapi
c 779	1.2	82237	50	AC003677	AC003677 Human P1	c 852	36	1.2	93273	32	AL135912	AL135912 Homo sapi
c 780	1.2	82426	76	AC055885	AC055885 Homo sapi	c 853	36	1.2	94106	11	AC005667	AC005667 Homo sapi
c 781	1.2	82690	31	AP001128	AP001128 Homo sapi	c 854	36	1.2	94148	41	HSJ858B6	AL118511 Homo sapi
c 782	1.2	83023	73	AC036231	AC036231 Homo sapi	c 855	36	1.2	94308	39	AC005480	AC005480 Homo sapi
c 783	1.2	83079	39	AC004538	AC004538 Homo sapi	c 856	36	1.2	94320	39	AC002564	AC002564 Human BAC
c 784	1.2	83514	31	AP001329	AP001329 Homo sapi	c 857	36	1.2	94320	39	AC002564	AC002564 Human BAC
c 785	1.2	83534	69	AC008933	AC008933 Homo sapi	c 858	36	1.2	94336	39	AC005250	AC005250 Homo sapi
c 786	1.2	84126	10	HS125N5	AL008728 Human DNA	c 859	36	1.2	94374	72	AC008384	AC008384 Homo sapi
c 787	1.2	84307	77	AC055742	AC055742 Homo sapi	c 860	36	1.2	94715	32	HSJ575F21	AL096841 Homo sapi
c 788	1.2	84412	10	HS973M2	AL033533 Human DNA	c 861	36	1.2	94960	39	AC000066	AC000066 Homo sapi
c 789	1.2	84464	10	HS782D21	AL023399 Human DNA	c 862	36	1.2	95038	72	AC011430	AC011430 Homo sapi
c 790	1.2	84633	39	AC005661	AC005661 citb_54_O	c 863	36	1.2	95267	10	HSJ117J24	AL024474 Human DNA
c 791	1.2	84871	10	HSJD54G6	AL121588 Human DNA	c 864	36	1.2	95367	39	AC004001	AC004001 Human BAC
c 792	1.2	84912	11	AF024533	AF024533 Homo sapi	c 865	36	1.2	95433	67	AC026437	AC026437 Homo sapi
c 793	1.2	85058	39	AC004989	AC004989 Homo sapi	c 866	36	1.2	95855	11	HSAC000115	AC000115 Human BAC
c 794	1.2	85058	72	AC011426	AC011426 Homo sapi	c 867	36	1.2	95956	39	AC006336	AC006336 Homo sapi
c 795	1.2	85335	11	HS136017	Z72001 Human DNA s	c 868	36	1.2	96256	10	HSJ117715	AL022315 Human DNA
c 796	1.2	85579	39	AC004744	AC004744 Homo sapi	c 869	36	1.2	96276	10	HSJ1175I6	AL049538 Human DNA
c 797	1.2	85654	10	HSP373C6	AL022393 Homo sapi	c 870	36	1.2	96308	39	AF190465	AF190465 Homo sapi
c 798	1.2	85863	72	AC010464	AC010464 Homo sapi	c 871	36	1.2	97060	55	AC010350	AC010350 Homo sapi
c 799	1.2	85931	12	AC007049	AC007049 Mus muscu	c 872	36	1.2	97392	10	HSBA150A6	AL096770 Human DNA
c 800	1.2	86049	80	AC0066585	AC0066585 Homo sapi	c 873	36	1.2	98117	11	AC003688	AC003688 Homo sapi
c 801	1.2	86394	32	HSJ84J3	AL117343 Homo sapi	c 874	36	1.2	98117	11	AC003688	AC003688 Homo sapi
c 802	1.2	86446	55	AC008754	AC008754 Homo sapi	c 875	36	1.2	98183	67	AC026448	AC026448 Homo sapi
c 803	1.2	86516	50	AC009493	AC009493 Homo sapi	c 876	36	1.2	98447	10	HSJ925J7	AL078622 Human DNA
c 804	1.2	86684	50	AC010605	AC010605 Homo sapi	c 877	36	1.2	98562	50	AC008969	AC008969 Homo sapi
c 805	1.2	86828	43	AC020569	AC020569 Homo sapi	c 878	36	1.2	98573	39	AF095725	AF095725 Homo sapi
c 806	1.2	86896	10	HS93N13	Z84489 Human DNA s	c 879	36	1.2	98684	32	AL133516	AL133516 Homo sapi
c 807	1.2	86969	79	AC063376	AC063376 Homo sapi	c 880	36	1.2	98697	39	AC004854	AC004854 Homo sapi
c 808	1.2	87013	10	HS528L19	Z84486 Human DNA s	c 881	36	1.2	98732	41	HSJ46801	AL109923 Homo sapi
c 809	1.2	87114	11	AC005353	AC005353 Homo sapi	c 882	36	1.2	98758	10	HSJ23H9	AL008582 Human DNA
c 810	1.2	87733	72	AC015600	AC015600 Homo sapi	c 883	36	1.2	98828	10	HSJ388E23	AL049552 Human DNA
c 811	1.2	88013	73	AC008402	AC008402 Homo sapi	c 884	36	1.2	99370	39	AC005057	AC005057 Homo sapi
c 812	1.2	88050	9	AP001421	AP001421 Homo sapi	c 885	36	1.2	99667	32	AL157376	AL157376 Homo sapi
c 813	1.2	88071	11	AC002316	AC002316 Homo sapi	c 886	36	1.2	99674	11	CNS01DSP	AL121838 Human chr
c 814	1.2	88133	74	AC026771	AC026771 Homo sapi	c 887	36	1.2	99902	10	HS59B16	AL032822 Human DNA
c 815	1.2	88215	39	AC006963	AC006963 Homo sapi	c 888	36	1.2	99935	32	AL136975	AL136975 Homo sapi
c 816	1.2	88217	41	AC0066407	AC0066407 Homo sapi	c 889	36	1.2	100000	9	AB020867	AB020867 Homo sapi
c 817	1.2	88882	10	HS448E20	Z97196 Human DNA s	c 890	36	1.2	100000	9	AP000011	AP000011 Homo sapi
c 818	1.2	89045	73	AC009185	AC009185 Homo sapi	c 891	36	1.2	100000	9	AP000013	AP000013 Homo sapi
c 819	1.2	89218	39	AC005902	AC005902 Homo sapi	c 892	36	1.2	100000	9	AP000018	AP000018 Homo sapi
c 820	1.2	89240	67	AC026420	AC026420 Homo sapi	c 893	36	1.2	100000	9	AP000020	AP000020 Homo sapi
c 821	1.2	89292	9	AP001252	AP001252 Homo sapi	c 894	36	1.2	100000	9	AP000030	AP000030 Homo sapi
c 822	1.2	89323	9	AP000021	AP000021 Homo sapi	c 895	36	1.2	100000	9	AP000033	AP000033 Homo sapi
c 823	1.2	89328	10	HS398C22	Z93784 Human DNA s	c 896	36	1.2	100000	9	AP000045	AP000045 Homo sapi
c 824	1.2	89800	69	AC010474	AC010474 Homo sapi	c 897	36	1.2	100000	9	AP000052	AP000052 Homo sapi
c 825	1.2	90082	11	AF045555	AF045555 Homo sapi	c 898	36	1.2	100000	9	AP000071	AP000071 Homo sapi
c 826	1.2	90220	10	HSJ734P14	AL049650 Human DNA	c 899	36	1.2	100000	9	AP000077	AP000077 Homo sapi
c 827	1.2	90299	11	HS72E17	AL033523 Human DNA	c 900	36	1.2	100000	9	AP000080	AP000080 Homo sapi

901	36	1.2	100000	9	AP0000081	AP0000081 Homo sapi	974	36	1.2	104070	42	AC018637_5	Continuation (6 of
902	36	1.2	100000	9	AP0000085	AP0000085 Homo sapi	c 975	36	1.2	104140	39	AC008122	AC008122 Homo sapi
c 903	36	1.2	100000	9	AP0000094	AP0000094 Homo sapi	c 976	36	1.2	104147	11	AC004703	AC004703 Homo sapi
c 904	36	1.2	100000	9	AP0000096	AP0000096 Homo sapi	c 977	36	1.2	104630	39	AC007848	AC007848 Homo sapi
c 905	36	1.2	100000	9	AP0000103	AP0000103 Homo sapi	c 978	36	1.2	104727	72	AC008632	AC008632 Homo sapi
c 906	36	1.2	100000	9	AP0000113	AP0000113 Homo sapi	c 979	36	1.2	104871	11	AC004584	AC004584 Homo sapi
907	36	1.2	100000	9	AP0000115	AP0000115 Homo sapi	980	36	1.2	105104	10	HS0347A17	AL049757 Human DNA
908	36	1.2	100000	9	AP0000120	AP0000120 Homo sapi	c 981	36	1.2	105127	50	AC004548	AC004548 Homo sapi
c 909	36	1.2	100000	9	AP0000128	AP0000128 Homo sapi	c 982	36	1.2	105138	39	AF124730	AF124730 Homo sapi
c 910	36	1.2	100000	9	AP0000133	AP0000133 Homo sapi	c 983	36	1.2	105284	69	AC010434	AC010434 Homo sapi
c 911	36	1.2	100000	9	AP0000133	AP0000133 Homo sapi	984	36	1.2	105334	10	HS914P20	AL034553 Human DNA
912	36	1.2	100000	9	AP0000137	AP0000137 Homo sapi	c 985	36	1.2	105345	41	HSJ272H18	AL109826 Homo sapi
913	36	1.2	100000	9	AP0000152	AP0000152 Homo sapi	c 986	36	1.2	105485	39	HUAC02301	AC002301 Homo sapi
914	36	1.2	100000	9	AP0000155	AP0000155 Homo sapi	987	36	1.2	105637	71	AP000782	AP000782 Homo sapi
c 915	36	1.2	100000	9	AP0000156	AP0000156 Homo sapi	988	36	1.2	105641	33	AC011483	AC011483 Homo sapi
c 916	36	1.2	100000	9	AP0000159	AP0000159 Homo sapi	c 989	36	1.2	105734	41	AC005000	AC005000 Homo sapi
c 917	36	1.2	100000	9	AP0000161	AP0000161 Homo sapi	c 990	36	1.2	105964	11	AL135940	AL135940 Human DNA
918	36	1.2	100000	9	AP0000167	AP0000167 Homo sapi	c 991	36	1.2	105964	11	AL135940	AL135940 Human DNA
c 919	36	1.2	100000	9	AP0000179	AP0000179 Homo sapi	992	36	1.2	105987	31	AP000617	AP000617 Homo sapi
c 920	36	1.2	100000	9	AP0000189	AP0000189 Homo sapi	c 993	36	1.2	106018	10	HS864I118	AL031293 Human DNA
921	36	1.2	100000	9	AP0000191	AP0000191 Homo sapi	994	36	1.2	106025	10	HS0404L14	AL024495 Human DNA
c 922	36	1.2	100000	9	AP0000198	AP0000198 Homo sapi	c 995	36	1.2	106039	39	AC004159	AC004159 Homo sapi
c 923	36	1.2	100000	9	AP0000200	AP0000200 Homo sapi	c 996	36	1.2	106102	70	AC027341	AC027341 Homo sapi
924	36	1.2	100000	9	AP0000201	AP0000201 Homo sapi	997	36	1.2	106373	10	HS1280I7	AL031654 Human DNA
c 925	36	1.2	100000	9	AP0000206	AP0000206 Homo sapi	998	36	1.2	106432	31	AP000860	AP000860 Homo sapi
926	36	1.2	100000	9	AP0000211	AP0000211 Homo sapi	c 999	36	1.2	106474	41	AC004977	AC004977 Homo sapi
c 927	36	1.2	100000	9	AP0000211	AP0000211 Homo sapi	c1000	36	1.2	106508	39	AC005102	AC005102 Homo sapi
c 928	36	1.2	100000	9	AP0000216	AP0000216 Homo sapi							
c 929	36	1.2	100000	10	AP0000508	AP0000508 Homo sapi							
c 930	36	1.2	100000	10	AP0000514	AP0000514 Homo sapi							
931	36	1.2	100000	10	AP0000517	AP0000517 Homo sapi							
c 932	36	1.2	100000	10	AP0000518	AP0000518 Homo sapi							
c 933	36	1.2	100000	10	AP0000519	AP0000519 Homo sapi							
934	36	1.2	100077	67	AC026170	AC026170 Homo sapi							
935	36	1.2	100184	50	AC008783	AC008783 Homo sapi							
936	36	1.2	100258	39	AC002462	AC002462 Human BAC							
937	36	1.2	100296	73	AC012324	AC012324 Homo sapi							
938	36	1.2	100701	39	AC005911	AC005911 Homo sapi							
939	36	1.2	100738	73	AC008662	AC008662 Homo sapi							
940	36	1.2	100849	10	HUMNEUROF	L05367 Human oligo							
941	36	1.2	100974	71	AC011461	AC011461 Homo sapi							
942	36	1.2	100997	10	HS262D12	Z99297 Homo sapien							
943	36	1.2	101002	32	AL136126	AL136126 Homo sapi							
c 944	36	1.2	101148	67	AC026471	AC026471 Homo sapi							
c 945	36	1.2	101196	10	HS463A9	280232 Human DNA s							
946	36	1.2	101270	10	HS483K16	AL034374 Human DNA							
947	36	1.2	101311	39	HUAC004513	AC004513 Homo sapi							
948	36	1.2	101368	72	AC016552	AC016552 Homo sapi							
c 949	36	1.2	101368	72	AC016552	AC016552 Homo sapi							
950	36	1.2	101484	55	AC011478	AC011478 Homo sapi							
c 951	36	1.2	101507	9	AP000359	AP000359 Homo sapi							
c 952	36	1.2	101685	11	HS039622	AL050341 Human DNA							
c 953	36	1.2	101685	40	AL133347	AL133347 Homo sapi							
954	36	1.2	101720	39	AC005006	AC005006 Homo sapi							
955	36	1.2	101846	10	HS197L1	AL031390 Human DNA							
956	36	1.2	101912	39	AC000100	AC000100 Homo sapi							
957	36	1.2	101952	10	AP001820	AP001820 Homo sapi							
c 958	36	1.2	101981	39	AC003001	AC003001 Homo sapi							
c 959	36	1.2	101999	40	CNS05TBT	AL163974 Homo sapi							
960	36	1.2	102019	11	HS012G14	AL078581 Human DNA							
961	36	1.2	102370	10	HSJ976013	AL117354 Human DNA							
c 962	36	1.2	102611	32	AL138734	AL138734 Homo sapi							
963	36	1.2	102634	42	AC018641_3	Continuation (4 of							
964	36	1.2	102663	75	AC034231	AC034231 Homo sapi							
c 965	36	1.2	102750	11	AC004465	AC004465 Homo sapi							
966	36	1.2	102818	11	AC003042	AC003042 Homo sapi							
c 967	36	1.2	102818	11	AC003042	AC003042 Homo sapi							
968	36	1.2	103146	10	HS187N21	Z98036 Human DNA s							
969	36	1.2	103367	10	HS0A22F16	AL049641 Human DNA							
970	36	1.2	103491	53	AC022432	AC022432 Homo sapi							
971	36	1.2	103530	39	AC006004	AC006004 Homo sapi							
972	36	1.2	103565	39	AC005022	AC005022 Homo sapi							
973	36	1.2	103746	31	AP0000769	AP0000769 Homo sapi							

ALIGNMENTS

RESULT 1

RNENACA	RNENACA	3117 bp	mRNA	08-MAR-1994
LOCUS	R. norvegicus mRNA for sodium channel, alpha subunit.			
DEFINITION	X70497			
ACCESSION	X70497.1	GI:458845		
VERSION	alpha subunit; sodium channel.			
KEYWORDS	Norway rat.			
SOURCE	Rattus norvegicus			
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;			
REFERENCE	1 (bases 1 to 3117)			
AUTHORS	Canessa,C.M., Horisberger,J.D. and Rossier,B.C.			
TITLE	Epithelial sodium channel related to proteins involved in			
JOURNAL	neurodegeneration			
MEDLINE	Nature 361 (6411), 467-470 (1993)			
REFERENCE	2 (bases 1 to 3117)			
AUTHORS	Canessa,C.M.			
TITLE	Direct Submmission			
JOURNAL	Submitted (02-MAR-1994) C.M. Canessa, Institut de Pharmacologie et			
FEATURES	Toxicologie, Bugnon 27, 1005 Lausanne, SWITZERLAND			
Source	Location/Qualifiers			
	1..3117			
	/organism="Rattus norvegicus"			
	/strain="Sprague Dawley"			
	/db_xref="taxon:10116"			
	/dev_stage="adult"			
	/tissue_type="distal colon epithelium"			
	82..2178			
	/gene="rENaC"			
	82..2178			
	/gene="rENaC"			
	/codon_start=1			
	/product="epithelial sodium channel, alpha subunit"			
	/protein_id="CAA49905.1"			
	/db_xref="GI:458846"			
	/translation="MLDHTRAPELINIDLDLHANSKPKSGMKGNFKQDPCFPQPMQK			
	/translaton="LQGDKREOGLGPEPSAPRPEEEALIEFHRSYRELQFQFCNNNTTHGAIKLVCS			
	KNRMKTAFAWLVLCFTFGMYQFALLFEYLSYPVSLNINLSNLSKLVFPFATVCTL			

NPYRTEIKEELEEDRIITEQTLFDLYKYNSSYTRQAGARRSSRDLGAFPHPLQRL
RTPPPYSORTARSSSSVDRNNPDQVDRKDKIGFOLCQNONKSDCFYQSYSSGVDAVR
EWRPHYINILSRLSDTSPALREELGNFIETCRFNOAPCNOANTSKFHPHMGNCYT
FNDKNSNLWMSMPGVNNGLSLTLRTEONDRIPLLSTVTGARVMVHGODEPAFMDDG
GFNLPGVETSLSMRKALDSLGGNYDCTENGSDVPKNLYPSKYTQOVCILHSCFEQ
NMIIKCGCAYIIFPKPGVEFCDYRKQSSWGYCYKLOGAFSLDSLUGFSKRPKCSV
INYLUSAGYSRMPKSDWIFEMISLQNNYITINKRNGVAKLNIFFKELNITKNS
PSVTWLSNLISGQSLWFGSVLSVEMADVIFDLVITLMLLRFRSRYWSPGR
GARGAREVASTPASSFPFRPHPTSPPLPQQGMTPLALTAPPAYATILGPSAPP
LDSAAPDCSACALAL"

BASE COUNT 684 a 983 c 753 g 697 t

ORIGIN

Query Match		100.0%;	Score 3117;	DB 12;	Length 3117;				
Best Local Similarity		100.0%;	Pred. No. 0;						
Matches 3117;		Conservative	0;	Mismatches	0;	Indels	0;		
				Gaps		0;			
QY	1	cegggtgcacccacgctgcgagggcagcctgggatgogggcagcgtcccgacagccc	60						
DB	1	CCGGGTGCACCCACGGCTCGAGGGCAGCGCTGGGATCGGGCAGCGTCCCGACAGCCC	60						
QY	61	catctgcctcaagctaataatgatctggaccacacagagccctgagctcaacattgac	120						
DB	61	CATTCTGCCCTTACGGTAATGATGCTGGACCAACAGAGCCCTGAGCTCAACATTGAC	120						
QY	121	ctagacctcagcctcccaactgcctaaagggctcattgaaggcaaccaattcaaggag	180						
DB	121	CTAGACCTTCAGCCCTCCAACTCGCCTAAGGGGTCCATGAAGGCCAACCAATTCAGGAG	180						
QY	181	caagaccttgtctcctcagcccatcgaaggactgggaagggggacaacagtgaagag	240						
DB	181	CAAGACCTTGTCTCTCAGCCCATGCAAGGACTGGGAAGGGGACAAACGTGAAGAG	240						
QY	241	caggcctggccggcgaacctcagcaccccgagcccaacagagagggagggcaactg	300						
DB	241	CAGGGCTGGGGCCGGAACCCCTCAGCACCCCGCAGCCCAACCGAGGAGGAGGCACTG	300						
QY	301	atgaattcaacgcctcctacgggagctctccagttcttcttgcaacaacacacacatc	360						
DB	301	ATTGAATTCACCGCTCTTACCGGGAGCTCTTCCAGTTCTTCTGCAACAACACCACTC	360						
QY	361	cacggggccatccgctgggtgtgtctcaaacacaaacccgeatgaagcggcctctggcg	420						
DB	361	CACGGGCCATCCGCTCGGTGTGTCTCCAAACAACAACCGCATGAAGCGGCTTCTGGCGG	420						
QY	421	gtcgtggtgtgacacctcgcatgatgtactggcagttcgcttgcgttcgaggag	480						
DB	421	GTGCTGTGGCTGTGCACCTTCGGCATGATGTACTGGCAGTTCCGCTTGTCTGAGGAG	480						
QY	481	tacctcagctacccagtgagcctcaacatcaacctcaattcagacaagctggcttccct	540						
DB	481	TACCTCAGCTTACCAGTGAGCCTCAACATCAACCTCAATTCAGACAACAGCTGTCTCCCT	540						
QY	541	gcgctcactgtctgcacccttaactctacagatacactgaaataaagaggagctggaa	600						
DB	541	GC CGTCACTGTCTGCACCTTAATCTTACAGATACACTGAATTAAGAGGAGCTGGAA	600						
QY	601	gagctgacgcatcagcagcagcagcgtttttgactgtacaaaatacaactcttctac	660						
DB	601	GAGCTTGACCGCATCAGGAGCAGAGCGCTTTTGACTTGTACAAATACAACTCTTCTAC	660						
QY	661	actgcagagctggggcccgacgcccagctcccgagacctctgggtgctttcccgac	720						
DB	661	ACTGCGCAGGCTGGGGCCCGACCGCGAGCTCCCGGACCTCTTGGGTGCTTCCCGCAC	720						
QY	721	ccctcagcgcctgcgactccacctccgacctactccggccgacgagcgcgagcggg	780						
DB	721	CCCTCGACGGCTGGGCACTCCACCTCCGCCCTACTCCGGCCGACGGCGCGAGCGGG	780						
QY	781	tcttcagcgtacgcagacaacatcccccaagtggaacggagagactggaagactcgcttc	840						
DB	781	TCTTCCAGCGTACGCGACACAATCCCAAGTGGACCGGAAGGACTTGAAGATCGGCTTC	840						

QY	841	caactgtgcaaccagaaataacagactgtttctaccagacatactcctctggggtggat	900
DB	841	CAACTGTGCAACAGCAAAATCAGACTGTTCTTACCAGACATACTCTCTGGGGTGGAT	900
QY	901	gcagtgaggagtggtaccgcttccatacatcaacattctgtccagactgtcgacacc	960
DB	901	GCAGTGGAGGTGGTACCGCTTCCATTATACATAAATCTGTGTCCAGACTGTTCGACACC	960
QY	961	tcgccgctctagagaagaagccctgggcaacttcattctcactgtcgtctcaaccag	1020
DB	961	TGCGCCCTCTAGAGGAAGAAGCCCTGGCAACTTCATCTTCACTGTGCGCTTCAACAG	1020
QY	1021	gccccctgcaaccaggcggaattattccaagttccaccaccccatgtacgggaactgctac	1080
DB	1021	GCCCCCTGCAACAGCGGAATATTTCAGTTCACCAACCCCATGTACGGGAACGTCTAC	1080
QY	1081	acttcaatgaagaagaactcactctcttgatgtctctcaatgcctgtgagtgacaacat	1140
DB	1081	ACTTCAATGACAAGAACAATCTCAATCTCTGGATGTCTCCATGCTGGAGTCAACAAT	1140
QY	1141	ggtttgtccctgacactgcacagacagaatgaacttcacccctgctgtccacagt	1200
DB	1141	GGTTGTCCCTGACACTGCGCACAGACGAATGACTTCATFCCCCCTGCTGTCCACAGTG	1200
QY	1201	acgggggcccagggtgatgggtcaggtcaggtcagcctgctcttatggatgagtggtgc	1260
DB	1201	ACGGGGCCAGGGTGTGGTGCATGGTTCAGGATGAGCGCTGCTTATGGATGATGGTGGC	1260
QY	1261	tcaacttgaggcctgggtggagacctccatcagtagagaaggagccctgggacagc	1320
DB	1261	TTCAACTTGAGGCGCTGGCGTGGAGACCTCCATCATAGTAGAAAGGAAGCCCTGGACAGC	1320
QY	1321	ctcggggaataattacggcgactgtactgagaatggttagcgatgtcccggttcaagaacct	1380
DB	1321	CTCGGAGGAATTACGGCGACTGTACTGAGAAAGGTGAGCGATGTCCCGGTCAAGAAGCTT	1380
QY	1381	taccttccaaagtatacacagcaggtgtgattcaactcctgcttccaggagaaacatgac	1440
DB	1381	TACCTTCCAAGTATACACAGCAGGTGTGCATTCACCTCTCTTCCAGGAGAACATGATC	1440
QY	1441	aagaagtgtgctgtgctacattcttaccctaaagcccaaggagggttggattctgtgac	1500
DB	1441	AAGAAGTGTGCTGTGCTTACATCTTACCTTAAGCCCAAGGAGTGTAGTTCTGTGTAC	1500
QY	1501	taccgaagcagagcctcctgggctattgctattataaaactgaaggcgctctcctctg	1560
DB	1501	TACCGAAAGCAGAGCTCCTGGGCTATTGTATTATTAAGTGCAGGGCGCTTCTCCTTG	1560
QY	1561	gacagcctgggctgtttctcctcaaggtgcggaagcctgtgtagtgatcaactacaasctc	1620
DB	1561	GACAGCCTGGGCTGTTCTCCAAGTGTGGAAGCCTTGTAGTGTGATCACTACAACTC	1620
QY	1621	cttgccggctactcactcaggtggccatctgtgaagtcaccagattggattcttcgagatctg	1680
DB	1621	CTGCGCGCTACTCACGGTGGCCATCTGTGAAGTCCCAAGGATTGGATCTTTCGAGATCTG	1680
QY	1681	tcottgcagaaacattacactattacaacaaagaacagaggttgcaagcgtcaaacatc	1740
DB	1681	TCCTTGAGAACAAATTACACTATTAAACAAGAAACAGGAGTTGCAAAAGCTCAACATC	1740
QY	1741	ttctcaagagactgaactataaaactaattcggagctcctctctgacagatggtcagc	1800
DB	1741	TTCTTCAAGGAGCTGAACATAAACTATAAAGTCTCCTTCTGTACAGATGGTCTCAGC	1800
QY	1801	ctcctgtccaaacctgggcagcagcagcctgtgggttggtcgtcgtgctctcgtg	1860
DB	1801	CTCCTGTCCAACTGGGCAGCCAGTGGAGCCTGTGGTTTGGCTGCTCCGTCTCTGTG	1860
QY	1861	gtggagatggcgacgtcatcttcgacctcctggttcacactctctcatgtgctacgc	1920
DB	1861	GTGGAGATGGCGAGCGTCATCTTCGACCTCTCTGGTCATCACACTTCTCATCTGCTCTACGC	1920

```
Qy 1921 cggttccggagcggtactgtctccaggacgagggccagggtgcccaggaggtggcc 1980
Db 1921 CGGTTCCGGAGCGGTACTGGTCTCCAGGAGAGGGGCCAGGGGTGCCAGGAGGTGCC 1980
Qy 1981 tccactcaggttctctcttcccggtcccggtttctgtctcactcactacatccccaccct 2040
Db 1981 TCCACTCAGGTTCTCTCTCCCGTCCCGTTCTGTCTCACCCTACATCCCCACCACCT 2040
Qy 2041 tctttgcccagcagcgatgacccctccccctggccctgacagccctcactgcctat 2100
Db 2041 TCTTTGCCCCAGCAGGCGATGACCCCTTCCCTGGCCCTGACAGCCCTCCACCTGCCAT 2100
Qy 2101 gctacttagccccagtgccctccactgactctcgccgctgactgtctgcctgt 2160
Db 2101 GCTACTTAGCCCCAGTGGCCCTCCACTGGACTCTCGGGCCCTGACTGTCTGCCGTGT 2160
Qy 2161 gccctggcgctctgagagagagagagatctctcaccagccctgagctcccttgt 2220
Db 2161 GCCCTGGCGGCTCTGAGAGAGAGAGAGATCTCTCACCCAGCCCTGAGCTCCCTTGT 2220
Qy 2221 aaactgatgatctcacttcagacacatcttccccagcgcccgcccgagccctctgtg 2280
Db 2221 AAAGTGTATGATATCTCATTACACATCTTCCCCAGCGGCCCGCCGAGCCCTCTGTG 2280
Qy 2281 gtgtcctgctgaggaacagcgctgagtaaggggccaggaagccgtccagagagcagg 2340
Db 2281 GTGTGCTGCTGAGGAGAACAGCGTGTAGTAAAGGGGCCAGGAAGCCGTCCAGAGCAGG 2340
Qy 2341 ggctaagatctctcagagctgacctgccccctgctctctgaaacactgccttccacaagca 2400
Db 2341 GGCTAAGATCTGCTCAGAGCTGCCCTGCCCTGCTTCTGAACACTGCTTCCACAAGCA 2400
Qy 2401 cagacaagtctctctccctcgatcagccagccagacttgagctctgacaaagca 2460
Db 2401 CAGACAAGTTCCCTCTTCCCTCGGATCAGCCAGCCAGACTGGAGTCTGACAAAGAAC 2460
Qy 2461 ttctctggagagacaaagcaataagaaagcatcgactcctacccccgtgacc 2520
Db 2461 TTTCTCTGGAGAGACCAAGGAACAATAGAACAGCCATCGCACTCTCTACCCCGTGACC 2520
Qy 2521 agagctggcctggcctcactgccttcaggagacactgatgtcgtcctcctctgaact 2580
Db 2521 AGAGGCTGGCGTGGCTCAGTCCCTTCAAGGACACTGATGTCTCTCTCTTGAAT 2580
Qy 2581 tgggtggggaaccccccaaaagccccctgttagttcttggcgaattcccccccctc 2640
Db 2581 TGGTGGGGAACCCCAAAAGCCCCCTTGTGTAGTCTTTTGGCAATCCCCCTCCCTC 2640
Qy 2641 actcctcaggtgggactagtagtaagacagacagcaggtaggcctggtgtctctctcat 2700
Db 2641 ACTCCTCAGGTTGGGACTAGAGTAAGACAGACCAGGGTAGGCCCTGGTGTCTCTCAT 2700
Qy 2701 tctcaagactctcattgtatcacaccataccccagtgacctctggtgtgtgtctcttc 2760
Db 2701 TCTCAAGACTCTCTATTGTATACACCATACCCAGTGCCTCTGTGGTGTGTCTCTCTC 2760
Qy 2761 tttctgttctcagcctggaagtcttctcactgtccctcagaacacacagctatgca 2820
Db 2761 TTTCTGTCTTCTCAGCTGGAGGTTCTTTTACATGTCTTCCCTCAGAACACAGCTATGCA 2820
Qy 2821 tcccttagaaacctgcccagaagcaccattcttctgtgaaactcagttaccctgttatgt 2880
Db 2821 TCCCTTAGAACCTTCCCCAAGGACCAATTCTTTTGTGAACCTCAGTTACCTGTTATGT 2880
Qy 2881 ctccccagaattgtgocctctccctcaccctcaccacagtgactgtattatgttcaac 2940
Db 2881 CTCCCCAGAATTGTGCGCCCTCTCCCTTCACCCCCACAGTGTACTGTATTATATGTTTCA 2940
Qy 2941 attcttttgtctgcctccctgggtaaacactgagctcctctgtgtggtcaggatgagatt 3000
Db 2941 ATTCTTTGTGTCGCTTCCCTGGGTAAACTGAGCTCCTTGTGTGGTCAGGGATGAGATT 3000
Qy 3001 tgcctgttttgtatccttcccggtctagcccgactctcccacttgggacgggtaggcaggt 3060
```

```
Db 3001 TGCCTGTTTGTATCTCTCCGGTCTAGCCAGTCTCCACTTGGGACGGGTAGSCAGT 3060
Qy 3061 actcaataaaggctgttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 3061 ACTCAATAAAGGCTTGTTCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3117

RESULT 2
RNASNAC 3081 bp mRNA ROD 07-DEC-1993
LOCUS R. norvegicus mRNA for amiloride sensitive Na+ channel protein.
DEFINITION X70521
ACCESSION X70521
VERSION X70521.1 GI:433909
KEYWORDS amiloride sensitive; Na+ channel protein.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 3081)
AUTHORS Lingueglia, E., Voilley, N., Waldmann, R., Lazdunski, M. and Barbry, P.
TITLE Expression cloning of an epithelial amiloride-sensitive Na+
channel. A new channel type with homologies to Caenorhabditis
elegans degenerins
JOURNAL FEBS Lett. 318 (1), 95-99 (1993)
MEDLINE 93170495
REFERENCE 2 (bases 1 to 3081)
AUTHORS Barbry, P.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-1993) P. Barbry, IPMC UPR411 CNRS, 660 route des
Lucioles 06560, Sophia Antipolis, Cedex, FRANCE
FEATURES
source
1..3081
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/dev_stage="adult"
/tissue_type="distal colon"
/cell_type="epithelial"
54..2153
/codon_start=1
/product="amiloride sensitive Na+ channel protein"
/protein_id="CAA49916.1"
/db_xref="GI:433910"
/db_xref="SWISS-PROT:P37089"
/translaton="MMLDHTRAPELINLIDLHASNSPKGSMKGNQKEQDPCPPQPMQ
GLGKDKREQGLGPEPSAPROPTEEELIEFHSYRELFFQFCNNTTHGAILRLVC
SKNRMKTAFWAVLWLCFTGMYWQFALLFEYLSPVSLNINLSNKLVPAYTVCT
LNPVRYTEIKEELELDRIETQFLDYKYNSYTRQAGARRRSRDLLGAFPHPLQR
LRTPPPYSGRTARSGSSSVRDNNPOVDKDWKIGFOLCQNKSDCFVQTVSSGVDAV
REWRPHYINILSRSDTSPALBEEALGNFTFCRFNQAPCNQANYSKFHHPTGNCY
TFNDKNSLWMSMPGVNGLSLTRTEQNDIFPLLSTVTGARVYHVGQDEPAMDQ
GFFNLKRGVETISMRKEALDSLGNNGDCTENGSDYFVKNLXPSKYTQQVCIHSCFQ
ENMKKGCAYIFYPKPKGVFCFDYRKQSSWGICYYYLQGLAFSLDGLGCFSKRPGS
VINYLKSGYSRWPVSKSDWIFEMLSLNNYTINNRKNGVAKLINERKELNYTNSF
SPSVTVSLILNLGSLWMLFGSSVLSSVEMAEILFDLLVITLLMLLRPRSRVYNSPG
RGARGAREVASTPASSPSPRCPTPTSPFSLPQQGWTPLALTAPPAYATLQPSAP
PLDGAAPDCSACALAL"
BASE COUNT 669 a 974 c 742 g 696 t
ORIGIN

Query Match 89.7%; Score 2796; DB 12; Length 3081;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3076; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 27 cagctggagtgaggacggtcccgagacgccccattctccttcacgcataatgatgt 86
Db 2 CAGCCTGGATGCGGGCAGCGTCCCGGACAGCCCCATTCTGCCTCAGCTAATGATGCT 61
Qy 87 ggaacacaccagagccccctgagctcaacattgaactagacctcaacgctccaaactgcc 146
Db 62 GGACACACAGAGCCCCCTGAGCTCAACATTGACCTAGACCTTACGGCTTACGGCTCCACACGCGC 121
```


|||||
Db 2282 AGTAAGGGGCCAGAGAGCGTCCAGAGGACAGGGGCTAACGATCTCTCAGAGCTGCC 2341
QY tgccctctgtctgaacactgccttc--cacaagcacagacaaagtccctctccctcgg 2424
|||||
Db 2342 TGCCCTGTCTTGAACACTGCCTTCCACACAGCAGACAGTCCCTCTTCCCTCGG 2401
QY atagccacagccagacttgagctctgacagaaacttctctggagagaccaaagaac 2484
|||||
Db 2402 ATCAGCCAGCCAGACTTGAGCTCTGACAGGAACCTTCTGGGAGAGGACCAAGAAC 2461
QY 2485 aaatagaacagcatcgactcactaccccgtagcagagagctggcctgcctcaactgcc 2544
|||||
Db 2462 AAATAGAACAGCCATCCACTCTACCCCGTGACCAAGAGCTGGCCTGCCTCACTGCC 2521
QY 2545 ttcaggagacactgatgtctctctctctctgaacttgggtggggaacccaccccaaaag 2604
Db 2522 TTcAGGAGACACTGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2581
QY 2605 cccctgttagttcttcttgcaatccctccctccctcactcactcaggtgggactagat 2664
Db 2582 CCCCTGTGTAGTCTTTGGCAATTCCCTCCCTCACTCTCTCAGGGTGGGACTAGAGT 2641
QY 2665 aagacagacagaggttaggctgtgtctctctctctctctctctctctctctctctctct 2724
Db 2642 AAGACAGACAGGAGTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2701
QY 2725 accatccccagtcgctctgtgtgtgtgtctctctctctctctctctctctctctctct 2784
Db 2702 ACNATACCCAGTGCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2761
QY 2785 ttcttcaactgtcctcagac 2844
Db 2762 TTCTTTCACTGTCCCTCAGAACACACAGTCTATGCATCCCTTAGAAACCTGCCAAGGC 2821
QY 2845 accattcttttgtaactcagttacccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2904
Db 2822 ACCATTTCTTTGTGAACTCAGTTACCCCTGTATGTCTCTCCCCAGAAATGTGCGCCCTC 2881
QY 2905 cccctcacccacagtgactgtattatattatattatattatattatattatattatatt 2964
Db 2882 CCTCACCCCAACAGTACTGTATTATATATGTTCACATTTCTTTGTGTGTGTGTGTGT 2941
QY 2965 gtaaacctgagctcctgt 3024
Db 2942 GTAAACTGAGCTCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3001
QY 3025 ctaggccagctctccacttgggacggttaggcaggttaggcaggttaggcaggttaggcag 3084
Db 3002 CTAGCCAGCTTCCACTTGGGAGGGGTAGGAGGTACTCAATAAAGGCTTGTTCATCA 3061
QY 3085 aaaaaaaaaaaaaaaaaa 3104
Db 3062 AAAAAAAAAAAAAAAAAA 3081

RESULT 3
RNU54700
LOCUS RNU54700 2200 bp mRNA ROD 14-JUN-1997
DEFINITION Rattus norvegicus epithelial sodium channel alpha subunit (rEnaca)
ACCESSION U54700
VERSION 54700
KEYWORDS U54700.1 GI:2148927
SOURCE Norway rat.
ORGANISM Rattus norvegicus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
AUTHORS 1 (bases 1 to 2200)
Kreutz, R., Struk, B., Rubattu, S., Hubner, N., Szpirer, J., Szpirer, C.,
Ganten, D. and Lindpaintner, K.
TITLE Role of the alpha-, beta-, and gamma-subunits of epithelial sodium
channel in a model of polygenic hypertension

JOURNAL Hypertension 29 (1), 131-136 (1997)
MEDLINE 97191134
REFERENCE 2 (bases 1 to 2200)
AUTHORS Kreutz, R.
TITLE Direct Submission
JOURNAL Submitted (11-APR-1996) Reinhold Kreutz, Medicine, Medicine, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA

FEATURES
source
1. .2200
/organism="Rattus norvegicus"
/strain="stroke-prone spontaneously hypertensive, Heidelberg (SHRSPHD)"
/db_xref="taxon:10116"
/tissue_type="kidney"
/chromosome="4"
/dev_stage="adult"
36. .2132
/gene="rEnaca"
36. .2132
/gene="rEnaca"
/codon_start=1
/product="epithelial sodium channel alpha subunit"
/protein_id="AAB61157.1"
/db_xref="GI:2148928"
/translation="MLDHTRAPELNLIDLHASNSPKGSKNQFKEQDPCPPMQG
LGKDRKEQGLGPEPSAPROPTREEEALIEFHRSYRELFOFFCNNTTHGAIRLVCS
KNRMKTAFWAVLWLCFTGMYWQFALLPEYLSYPVSLNINLSDKLVPPAVTCTL
NPYRTEIKEELEELDRITQTLFDLYKYNSSYTRQAGARRSRDGLGAPPLQRL
RTPPPYSGRTARSGSSVRDNDPQVDRDKWKIGFQLCNKNKSDCFOTYSSGVDAVR
EYFRFYINILSRLSDTSPALEEALGNFTFCFNOAPCNQARYSKFHHPMGNCYT
FNDKNSNLMMSSMPGVNNGLSLTLEQNDIFPLISTVTGARMVHQDEPAFMDG
GNLRPGVETSISMRKALDSAGNYGDCDENGSDYKVKLYPSKYTOQVCIHSCFOE
NMLTKGCVAIFYPKPGVEFCYRKQSSWGYCYKLOGAFSLDSGCSKCRKQCSV
INYLKAGSRWFSVKSQDWIFEMLSLQNNYTNKRNKGVAKLINIFKELNYTNSES
PSVTWLSLNSGQSLWFSVSVVENAELIFDLVITLLMLLRFRKRYWSPGR
GARGREVASPTASSPFRFCPHPTSPPLPQQGMPPLALTPPPAYATLGPSPAPP
LDSAATPDCSACALAAI"
BASE COUNT 480 a 699 c 545 g 476 t
ORIGIN

Query Match 67.3%; Score 2098; DB 12; Length 2200;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 47 gtcccgacagccattctgccttcacgctaataatgactgtgacacacacagaccctg 106
Db 1 GTCCCGGACAGCCCATTCCTGCTTACGCTTAATGATGCTGGACACACAGACCCCTG 60

QY 107 agctcaacattgacctagacctcacctccactcgcctaaagggtcccatgaaggca 166
Db 61 AGCTCAACATTGACTAGACCTTCACGCTTCACTCGCTTAAGGGTCCATGAAGGCA 120
QY 167 accaattcaaggagcaagacccttgcctcctcagcccatcaggaagactggggagggg 226
Db 121 ACCAATTCAAGGAGCAAGACCCCTTGTCTCTCAGCCCATCAAGAGACTGGGAGGGG 180
QY 227 acaaacgtgaagacagagccctggcccggaacccctcagcaaccccgagccaccagag 286
Db 181 ACAAACTGAAGACAGAGGGCCCTGGGCGGAAACCTTCAGCACCCCGGACCCACCGAGG 240
QY 287 agagaggagcactgatgaattccaccgctcctaccgggagctcttcagttctctgca 346
Db 241 AGGAGGAGGACTGATTGAATTCACCGCTCTTACCGGGAGCTCTTCAGTTCCTCTGCA 300
QY 347 acaaacacacacacacaggggcccacccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 406
Db 301 ACAACACACACATCCACGGGGCCATCCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 407 cggcctctctggcggtgctgtgctgtgcaccttcggcagctatgtactgagcttcgct 466
Db 361 CGGCCCTTCTGGCGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420

Qy	467	tgctgttcgaggagtagtacctcagctacccagtgagcctcaacatcaacctcaattcagaca	526
Db	421	TGCTGTTCGAGGAGTAGTACCTCAGTACCCAGTACGAGCTCAACATCAACCTCAATTTCAGACA	480
Qy	527	agctggcttcctccctgcscgtcactgctcagcccttaactcctacagatacacactgaattta	586
Db	481	AGCTGTGCTTCCTCGCGTCACTGTCTGCACCTTAAATCCCTTACAGATACACTGAATTA	540
Qy	587	aagagagctggaagagctggaccgcacacagagcagacgccttttggactgtacaaaat	646
Db	541	AAGAGGAGCTGGAAGAGCTGGACCGCATCACGGAGCAGACGCTTTTGTGACTGTACAAAT	600
Qy	647	acaaacttctcaactcgtccagagctggggcccgagcgcgcagctcccgagacctcctgg	706
Db	601	ACAACCTCTTCCTTACATTCGCCAGCTGGGGCCGACGCCGACGTCCCGGACCTCGTGG	660
Qy	707	gtgctttccgcagccccctcgagcgtcgcactccacctccgcctactccggcgccga	766
Db	661	GTGCTTTCGCGCAGCCCTTGACGGCGCTCGGCACATCCACTCCGCCCTACTCCGGCGCGCA	720
Qy	767	cgcgcgagcggggtttccagcgtacgcgacaaatccccaagtggacgcggaagact	826
Db	721	CGCGCGCAGCGGCTTCTCAGCGTACGGACAACAATCCCAAGTGGACCGGGAAGGACT	780
Qy	827	ggaagatcgggtctccaactgtgcaaccagaaacaaatcagactgtttctaccagacatact	886
Db	781	GGAAGATCGGCTTCCAACTGTGCAACCAGAACAAATCAGACTGTTTCTACACAGACATACT	840
Qy	887	ccctcggggtggatgcagtgagggagtggtaccgcttccattacatacaacattctgccca	946
Db	841	CTCTCGGGTGGATGCAGTGAGGGAGTGTACCGCTTCCATTTACATCAACATTCCTGTCCA	900
Qy	947	gaactcggacacctgcgcgtctagaggaagaagccctgggcaacttcatcttcacct	1006
Db	901	GACTGTCCGACACTCGCCGCGTCTAGAGGAAGAAGCCCTGGGCAACTTCATCTTCACCT	960
Qy	1007	gtcgcttcaacaggccccctcaaccaggcgaattattccaagtttccacaccccactgt	1066
Db	961	GTGCTTCAACAGGCCCTCTGCAACGAGCGGAATATTTCGAAGTTCCACCACCCTCATGT	1020
Qy	1067	acgggaactgtcacacttticaatgacaagaacaaactccaaatctctggatgtcctccatgc	1126
Db	1021	ACGGGAATGCTACACTTTCATATGACAAGAACAACCTCCAATCTCTGGATGTCTCCATGC	1080
Qy	1127	ctggagtcacaatggtttgtccctgaaactgcgcaagcagacagaatgaattcatcccc	1186
Db	1081	CTGGAGTCAACAATGGTTGTCCCTGACACTCGCGACAGAGCAGAATGACTTCATCCGCC	1140
Qy	1187	tgctgtccacagtgacggggccaggtgatggtgcattggtcagagtagagcctgccttta	1246
Db	1141	TGCTGTCCACAGTGACGGGGCCAGGGTGATGGTGCATGGTACAGTAGAGCTGCCCTTTA	1200
Qy	1247	tgatgatggtggtcttcaacttgagcctggcgtggagaccttccatcatgatgagaagg	1306
Db	1201	TGGATGATGTTGCTTCACTTGAGSCCTGGCGTGGAGACCTCCATCATGATGAGAAAGG	1260
Qy	1307	aagccctggacagcctcggagaaaattacggcgactgtactgagaatggttagcagtcc	1366
Db	1261	AAGCCCTGGACAGCCTCGGAGGAAATTAACGGCGAGCTACTGAGAATGGTAGAGCATGTCC	1320
Qy	1367	cggtcagaacacctttacccttccaagtatacacacagcaggtgatcattcactcctgtcc	1426
Db	1321	CGGTCAAGAACCTTTACCTTCTCAAGATATACACAGCAGGTGTGCATTCACGTCTCTTC	1380
Qy	1427	aggagaacatgatacgaagaagtgtgctgtgcctacattcttaccctaagccccagggag	1486
Db	1381	AGGAGAACATGATCAAGAAAGTGTGGCTGTGCCCTACATCTTCTACCTTAAGCCCCAAGGAG	1440
Qy	1487	ttgagtctgtgactaccgaagcagagcctcctggggcta ttgctattataaactgcagg	1546
Db	1441	TTGAGTCTGTGACTACCGAAAGCAGAGCTCCTGGGGTATTGGTATATATAAATCGCAGG	1500

QY	1547	g	gcgccttctctcttgacagccctgggctgtttctccaaagtgcggaagcccttgtagtctga	1506
Db	1501	G	GGCCCTTCTCCTTGGACAGCCCTGGGCTGTTCTTCCAAAGTCTCGAAGACCCCTTGCTAGTCTGA	1560
QY	1607	t	caactcaaaactctctccggctactcacggtggccatctgtgaagtcocagattgga	1666
Db	1561	T	CAACTACAAACTCTCTGCCGGCTACTCAGGTGGCCATCTGTGAAGTCCAGGATTGGGA	1620
QY	1667	t	cttcgagatgctgctctgcagaaacaattacactattaaacaacaaagaaacggagcttg	1726
Db	1621	T	CTTCGAGATGCTGCTTTCGAGAACAAATTACACTATTACACAAACAAACGAGTTG	1680
QY	1727	c	aaagctcaacatctcttctcaaggagctgaactataaaactaattcggagctctctctctg	1786
Db	1681	C	AAAGCTCAACACTCTTCTTCAAGGAGCTGAACATATAAAACTATAATTCGGAGTCTCCTTCTG	1740
QY	1787	t	caacatggtcagctctctgtccaaactctgggaacagtcagtcgagcctgtggttggctcgt	1846
Db	1741	T	CACGATGCTGAGCTCCCTGTCCAACTGGGACGAGTGGAGCCGTGTGGTTTGGCTTCGT	1800
QY	1847	c	cgctgctctctggtggagatggcggaagtcactctcgacctctctggtcatcacacttc	1906
Db	1801	C	CGTGCTCTCTGTGGTGGAGATGGCGGAGCTCATCTTCGACCTCCTGGTCAATCACACTTC	1860
QY	1907	t	aatgctgtacgcgcggttcctggagcgcgtactggtctccagagacagggccaggggtg	1966
Db	1861	T	CATGCTGTACGCCGGTTCCTGGAGCCGCTACTGGTCTCCAGGACGAGGGCCAGGGGTG	1920
QY	1967	c	caaggaggctgcctccactcacagttctctctctccgtctccgtctctctcacccta	2026
Db	1921	C	CAAGGAGGTGGCCCTCCACTCAGACTTCCTCTCTCCCTCCCGTTCTGTCTCCACCCCTA	1980
QY	2027	c	atccccaccactctcttggccccagcagggatgaccctcccttggccctgacagccc	2086
Db	1981	C	ATCCCCACCACCTCTTTGGCCCCAGAGGGATGACCCCTCCCTCGGCCCTGACAGCCC	2040
QY	2087	t	ctccacctgcctatgctactctagggcccgagtcgcccctccactcgcgactgcgcgcgtg	2146
Db	2041	T	CTCCACCTGCCCTATGCTACTCTAGGCCCCAGTGCCTCTCCACTGGACTCTGCGGCGCTG	2100
QY	2147	a	ctgttctgctgtcccttggcgcgctctgagagaggagaagatcctctcaccagcc	2206
Db	2101	A	CTGTTCTGCCCTGCCCCGCGCGCTCTGAGAGAGGAGAGGATCCTCTCACCAGCC	2160
QY	2207	ct	gagctccctctgtaaacctgattgatctcactctcagca	2246
Db	2161	CT	GAGCTCCCTTGTAACATGATGGATATCTCACTTCCAGCA	2200
RESULT 4				
RNU54699				
LOCUS	RNU54699	2200 bp	mRNA	ROD
DEFINITION	Rattus norvegicus epithelial sodium channel alpha subunit (rEnaca)			
ACCESSION	U54699			
VERSION	U54699.1			
KEYWORDS	GI:2148925			
SOURCE	Norway rat.			
ORGANISM	Rattus norvegicus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
AUTHORS	1 (bases 1 to 2200) Kreutz, R., Struk, B., Rubattu, S., Hubner, N., Szpirer, J., Szpirer, C., Ganten, D., and Lindpaintner, K.			
TITLE	Role of the alpha-, beta-, and gamma-subunits of epithelial sodium channel in a model of polygenic hypertension			
JOURNAL	Hypertension 29 (1), 131-136 (1997)			
MEDLINE	97191134			
REFERENCE	2 (bases 1 to 2200)			
AUTHORS	Kreutz, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (11-APR-1996) Reinhold Kreutz, Medicine, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115.			

RESULT	4
RNU54699	
LOCUS	2200 bp mRNA
DEFINITION	Rattus norvegicus epithelial sodium channel alpha subunit (rEnaca) mRNA, complete cds.
ACCESSION	U54699
VERSION	U54699.1 GI:2148925
KEYWORDS	.
SOURCE	Norway rat. Rattus norvegicus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 2200)
AUTHORS	Kreutz,R., Strub,B., Rubattu,S., Hubner,N., Szpirer,J., Szpirer,C. Ganten,D. and Lindpaintner K. Role of the alpha-, beta-, and gamma-subunits of epithelial sodium channel in a model of polygenic hypertension Hypertension 29 (1), 131-136 (1997) 97191134
TITLE	Hypertension 29 (1), 131-136 (1997)
JOURNAL	
MEDLINE	
REFERENCE	2 (bases 1 to 2200)
AUTHORS	Kreutz,R.
TITLE	Direct Submission
JOURNAL	Submitted (11-APR-1996) Reinhold Kreutz, Medicine, Medicine, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115.

FEATURES	USA	Location/Qualifiers	QY	587	aagagagctggaagagctggaccgcatcagcgagcagcgtcttttgactgtgacaaat	646
source		1. .2200	Db	541	AAAGAGAGCTGGAGAGCTGGACCGCATACGGAGCAGACGCTTTTGTACTGTGACAAAT	600
		/organism="Rattus norvegicus"	QY	647	acaactctctactactcgcagcagctggggccgcgacccgcgacgtcccgacactctg	706
		/db_xref="taxon:10116"	Db	601	ACAACTCTTCTACACTCGCCAGGCTGGGGCCGACGCCGACGCTCCCGGACCTCTCTGG	660
		/dev_stage="adult"	QY	707	gtctttcccgacccccctcagcgctcgcgactccacactcccgccctactccggcgcca	766
gene		/issue_type="kidney"	Db	661	GTGCTTTCCGCGCACCCCTGCTGCGCTGGCGACTCCACCTCCGCGCTACTCCGCGCGCA	720
CDS		/chromosome="4"	QY	767	cgcgcgcagcgggtcttcagcgctacgcgacacacacacacacacacacacacacacac	826
		36..2132	Db	721	CGGCGCGCAGCGGCTCTTCAGCGTACGCGACACAAATCCCAAGTGGAGGAGACT	780
		/gene="rEnaca"	QY	827	ggaagatcggtcttccaaactgtcaaccacgaacaaacagactgtttctaccagacatact	886
		/codon_start=1	Db	781	GGAAGATCGGCTTCCAACTGTGCAACACAGAAACAAATCAGACTGTTTCTACCAACATAC	840
		/product="epithelial sodium channel alpha subunit"	QY	887	ctctgggggtggatgcaatgagggagtggtaccgcttccattaccatacaaatctgtcca	946
		/protein_id="AB61156.1"	Db	841	CCTCTGGGGTGGATGCGAGTGAGGAGTGGTACCCTTCCATTACATCAACATTCGTCCA	900
		/db_xref="GI:2148926"	QY	947	gactgtcgacacactcgcgcgtctagaggaagccctgggcaacttcatcttccact	1006
		/translation="MLDHTRAPELNIIDLHASNPSKSMKNQFKEODPCPPQPMOG	Db	901	GACTGTGGACACCTCGCCGCTCTAGAGGAGAGCCCTGGGCAACTTCATCTTCACCT	960
		LGGKREBQGLGPEAPRQTEEEALIEFHSRYELFPFCNNNTTIGHAIRLVCS	QY	1007	gtcgcttcaaccaggccccctgcaaccaggcgaattatttccaaagtccacacccccatgt	1066
		KHNKMTAFWALWLTFCGMWQFALLFEVLSYVSLNINLSKLVFPVAVTCTL	Db	961	GTGCTTTCAACAGCGCCCTTGCAACACAGCGGAATTTATTCAGTTCCACACCCCATGT	1020
		NPYRTEIKEELELDRIEOTLFDLYKYNSSYTRQAGARRSRDLGAFPLQRL	QY	1067	acgggaaactgctacatttcaatgacaagaacaaactccaatctctggatgctctccatgc	1126
		RPPPPYSRTARSQSSYRDNNPQVDKDKIGFQLCNQKSDCEFYQYISGVDAVR	Db	1021	ACGGAACTGCTACACTTTCAATGACAAGAACAACTCCAATCTCTGTGATGCTCTCCATGC	1080
		EWRFYINILSRSDTSALBEEALGNFICTRFQAPCNQANYSKFFHPMTGNCT	QY	1127	ctgagctcaacaatggtttgctcctgacactgcgcacagacagaatgactctcctcccc	1186
		FNDKNSLWMSMPGVNGLSLTLTEQNDIPLSTVTGARVMVHGDEPAFMDGG	Db	1081	CTGGAGTCAACATGGTTTGTCCCTGACACTGCGCACAGACAGACAGTTCATTCCTCCC	1140
		GNLRGVETSISMREALDSLGNYGCTENGSDVPVKLYPSKYTQOVCIHSCFOE	QY	1187	tgctgtccacagtgacggggggcagggtgagtggtgcatggtcaggtgagcctgctctta	1246
		NNIKKGCAYIFPKPKGVFCDYRKOSWGYCYKLOGAFSLDSGCFCKRKPCSV	Db	1141	TGCTGTCCACAGTACAGGGGGCCAGGCTGATGTGTGTCATGGTGCAGATCAGCTGCCTTTA	1200
		INVKLSAGYSRWPVKSDWIFEMLSLNNYTIINNRGVAKLNIFFKELNKTNSE	QY	1247	tgatgagtggtggttcaacttgaggcctgggtggagacctccatcagtagagaaagg	1306
		PSVTWVLSLNLGWSLWFGSSVLSVEMAEILFDLLVITLLMLLRFRSRYSWSPGR	Db	1201	TGGATGATGTGTGGCTTCAACTTGGAGGCTTGCGCTGGAGACCTCCATCAGTATGAGAAGG	1260
		GARGAREVATPASSPFRCPHTSPSPSLPQQGMTPPLAULTAPPATATLGPSPAP	QY	1307	aagcccttgacagcctcgaggaaattacgcgactgtactgagaatggttagcagtgctc	1366
		LDSAPDCSACALAL"	Db	1261	AGCCCTTGGACAGCTCGGAGGAAATTCAGCGGACTGTACTGAGANTGGTAGCATGTCC	1320
			QY	1367	cggtcagaagaacctttacccttccaaagtatacacagcaggtgtgcatcactcctgcttcc	1426
			Db	1321	CGGTCAAGAACTTTACCTTTCCAAAGTATACACAGCAGGTGTGCATTCACCTGCTTCC	1380
			QY	1427	aggagaacatgatcaagaagtgtggtgtgacctacatcttccacttaagccccaggag	1486
			Db	1381	AGGAGAACATGATCAAGAAAGTGTGGCTGTGCTTACCTTACCCCTAAGCCCCAAGGAG	1440
			QY	1487	ttgagttctgactaccgaaagcagagctcctgggctattgtattataaaactgcagg	1546
			Db	1441	TTGAGTCTGTGACTACCGAAGCAGAGCTCCTGGGCTATTGCTATTATTAACATGACAGG	1500
			QY	1547	gcgccttctccttgagcagcctgggctgttcttccaaagtgcgaaagccttctgtagtga	1606
			Db	1501	GGGCTTCTCCTTGGACAGCCTGGGCTGTTCCTCCAAAGTGTGCGAAGCCTTGTAGTGTGA	1560
			QY	1607	tcaactcaaaactctctgcgggtactcaacggtggccatctgtgaatccccaggattgga	1666
			Db	1561	TCAACTAAACTCTCTGCCGCTACTCAGCGTGGCCATCTGTGAAGTCCCGAGGATGGA	1620
BASE COUNT	479 a	699 c	546 g	476 t		
ORIGIN						
Query Match 65.7%; Score 2047; DB 12; Length 2200;						
Best Local Similarity 99.9%; Pred. No. 0;						
Matches 2197; Conservative 0; Mismatches 3; Indels 0; Gaps 0;						
QY	47	gtccggagacccattcttcacgctaatgatgctggaccacacagagccctg	106			
Db	1	GTCCCGACAGCCCATCTGCTTACGCTTAATGATGCTGGACACACAGAGCCCTG	60			
QY	107	agctcaacattgactagaccttcaacgctccactgcctcaaggggtccatgaaggcca	166			
Db	61	AGCTCAACATTGACTTAGACTTCACGCTCCAACTCGCTCAAGGGTCTCATGAAGGCA	120			
QY	167	acaaattcaaggagacacccctgtctcctcactcagccatgcagggagggagggg	226			
Db	121	ACCAATTCAGGAGCAGAGACCTTGTCTCTCAGCCCATGCAAGGACTGGGGAAGGGG	180			
QY	227	acaaactgaagcagggcctggcccggaacctcagcaccccgacccacccaggg	286			
Db	181	ACAAACGTGAAGACAGGGCTGGGCCCGGAACCTCAGCACCCCGCAGCCACCGAGG	240			
QY	287	aggaggagcactgattgaattccacgcctcctaccgggagcttccagttctcttgc	346			
Db	241	AGGAGGAGCAGTATGATGAGTTCACCGCTCCTACCGGGAGCTTCCAGCTTCTTCTGCA	300			
QY	347	acaac	406			
Db	301	ACAACACACACATCCAGGGGCCATCCGCGCTGTGTGCTGCCAAACACACACGATGA	360			
QY	407	cggccttctggcggtgtgtggtgtgaccttcgcatgatgtactgagcttgcct	466			
Db	361	CGGCTTCTGGCGGGTGTGTGGCTGTGTGACCTTCGGCATGATGTACTGGCAGTTCGCT	420			
QY	467	tgctgttcgaggagactcagctaccagctgagcctcacaacataacactcaatccagaca	526			
Db	421	TGCTGTTGAGGAGTACTCAGTACCCAGTGAGCTCAACATCAACCTCAATTCACACA	480			
QY	527	agctgttcttccctgcgctcactgtctgcacccttaactcattacagatacactgaatta	586			
Db	481	AGCTGGTCTTCCCTGCGGCTACTGCTGCACCCCTTAATCTCTTACAGATACACTGAAATTA	540			


```
mRNA      1931. .>2223
/genes="ENACA"
/notes="alternative transcription start site"
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
1993. .>2223
/genes="ENACA"
/notes="ion channel"
/codon_start=1
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
/protein_id="AAB16017.1"
/db_xref="GI:4322329"
/translation="MMLDHTRAPELNLIDLHASNSPKGSMKGNQKPEDPCPPQMQ
GLGKGKREOGLGPPSAPROPTEEALIEP"
BASE COUNT      432 a      673 c      639 g      479 t
ORIGIN

Query Match      9.1%; Score 283; DB 12; Length 2223;
Best Local Similarity 100.0%; Pred. No. 9.9e-156;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 cagcctggatgaggcagcgtcccgacagcccccattctgccttcacgtatgatgct 86
|||||
Db 1941 CAGCCTGGATGCGGACAGCGTCCCGACAGCCCCATTCGCCCTCACGGCTAATGATGCT 2000
|||||

QY 87 ggaccacacagagccctcagctcaacattgaaccttcaacgtccctcccaactcgcc 146
|||||
Db 2001 GGACACACACAGAGCCCTCAGCTCAACATGTACCTAGACCTTCACGCCCTCCAACTCGCC 2060
|||||

QY 147 taagggtccatgaaggcgaacaaattcaaggagcagccctgtcctcagcccat 206
|||||
Db 2061 TAAGGGTCCATGAAGGCGACCAACCAATTCAGGACCAAGACCCTTGCTCCTCAGCCCAT 2120
|||||

QY 207 gcaaggactgggaagggggacaaactgaagagcagggcctggcccggaacctccagc 266
|||||
Db 2121 GCAAGGACTGGGAAGGGGACAAACCTGAAGACAGCGGCTGGGCCCGGAACCTTCAGC 2180
|||||

QY 267 accccggcagccacacagagagaggaggagggcactgattgaattc 309
|||||
Db 2181 ACCCCGACGCCACCAGGAGGAGGAGGAGGCACTGATTGAATTC 2223
|||||

RESULT 7
RNCCK3UTR      252 bp      mRNA      ROD      04-APR-1997
DEFINITION      R.norvegicus mRNA for 3'UTR of amiloride sensitive sodium channel
protein.
ACCESSION      X89818
VERSION      X89818.1 GI:1419186
KEYWORDS      3' UTR; amiloride sensitive sodium channel protein; CDK3 protein.
SOURCE      Norway rat.
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE      1 (bases 1 to 252)
AUTHORS      Page,R., Morris,C., Williams,J., von Ruhland,C. and Malik,A.N.
TITLE      Isolation of diabetes-associated kidney genes using differential
display
JOURNAL      Biochem. Biophys. Res. Commun. 232 (1), 49-53 (1997)
MEDLINE      97236278
REFERENCE      2 (bases 1 to 252)
AUTHORS      Page,R.A.
TITLE      Direct Submission
JOURNAL      Submitted (18-JUL-1995) R.A. Page, University of Wales College of
Cardiff, School of Molecular and Med.Biosciences, P.O. Box 911,
Cardiff CF1 3US, UK
COMMENT      Related sequence X70521.
FEATURES
source      Location/Qualifiers
1. .252
/organism="Rattus norvegicus"
/strain="Goto-Kakizaki (GK) Wistar"
```

```
/db_xref="taxon:10116"
/dev_stage="adult"
/tissue_type="kidney"
/clone="CDK3 (pSEM clone)"
1. .252
/genes="CDK3"
3'UTR      1. .252
/genes="CDK3"
polyA_signal 221..226
/notes="amiloride sensitive sodium channel protein"
/genes="CDK3"
polyA_site 231
/genes="CDK3"
BASE COUNT      48 a      77 c      51 g      76 t
ORIGIN

Query Match      4.5%; Score 141; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.3e-72;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2905 ccctcacccccacagtgactgtattatattatgttcacattctttgtctcctccctgg 2964
|||||
Db 62 CCCTCACCCCCACAGTGTACTGTATTATATGTTTCACATTCTTTGTGTCTCCCTCCCTGG 121
|||||

QY 2965 gtaaacctgagctcctgtgtggtcagggatgagattgctctgtttgtatccttcgggt 3024
|||||
Db 122 GTAAACCTGAGCTCCTGTGTGTCAGGGATGAGATTGCTCTGTTTGTATCTCTCCGGT 181
|||||

QY 3025 ctagcccagctcccaacttgg 3045
|||||
Db 182 CTAGCCAGTCTCCCACTGG 202
|||||

RESULT 8
AF112185      3000 bp      mRNA      ROD      27-MAR-2000
LOCUS      Mus musculus epithelial sodium channel alpha subunit mRNA, complete
DEFINITION      cds.
ACCESSION      AF112185
VERSION      AF112185.1 GI:4469398
KEYWORDS      house mouse.
SOURCE      Mus musculus
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 3000)
AUTHORS      Ahn,Y.J., Brooker,D.R., Kosari,F., Harte,B.J., Li,J., Mackler,S.A.
and Kleyman,T.R.
TITLE      Cloning and functional expression of the mouse epithelial sodium
channel
JOURNAL      Am. J. Physiol. 277 (1), F121-F129 (1999)
MEDLINE      99345762
REFERENCE      2 (bases 1 to 3000)
AUTHORS      Ahn,Y.J., Brooker,D.B. and Kleyman,T.R.
TITLE      Direct Submission
JOURNAL      Submitted (09-DEC-1998) Renal Electrolyte Division, University of
Pennsylvania, 422 Curie Boulevard, 700 Clinical Research Building,
Philadelphia, PA 19104, USA
FEATURES
source      Location/Qualifiers
1. .3000
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="6"
/map="distal"
32. .2131
/codon_start=1
/product="epithelial sodium channel alpha subunit"
/protein_id="AAB21244.1"
/db_xref="GI:4469398"
/translation="MLDHTRAPELNLIDLHASNSPKGSMKGNKFNKEDLCPPPLPMQG
LGKDKREQALGPPEPQPTPEEALIEFHSRYRELFFQFNNTTIHGAIRLVC
```

KHKMTAFMAVLWLCTFGMTYQFALLFEYFSPVSLNINLSKDLVFPVAVTVCTL
NPYRYTEKLEEDLRIPEQTLFDLYKYNSTYQAGGRRSTRDLRGALPHQLRL
RTPPPNPARSASSSVDRNNDPOVDKDWKIGFOLCNOXSDCFYQYSSGVDAVR
EWYRHYINILSRLPDTSPALKEEALGSFICTCRNOAPCNOANTSOEHPHXYNCYT
FNKNLNLMWSMGVYNGLSLTLTEONDPIPLSTVTGARVWVHQDEPAFMDGG
GNVRPGVETISIMRKEALDSLGGNNGDCTENGSDVPVKNLYPSKYTVQOVCIHSCFQE
NNIKKGCAGIPIYPRKGVFECDYLKQSSWGYYKLOAFSLDSLGCFSKRKPQSV
TNYKLISAGYSRWPVSQSDWIFEMLSLQNNYTIINNRNGVAKLINIFFKELNKTNSCS
PSVTWVSLNLSGOWSLWFGSSVLSVEMAEILFDLLVITLIMLLHRRFSRWSPGR
GARGAREVASTPASSPSPRCPHPTSPPSLSLQQGTTPPLALTAPPAYATLGPSP
LDSAVPGSSACAPANAL"

BASE COUNT 666 a 958 c 701 g 675 t
ORIGIN

Query Match 3.3%; Score 104; DB 12; Length 3000;
Best Local Similarity 100.0%; Pred. No. 9.6e-50;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1120 tccatgcctggagtcacaaatggtttgctccctgacactgcgcacagcagaatgacttc 1179
|||||

Db 1070 TCCATGCCTGGACTCAACATGTTGTCCCTGACACTGCGCAGACAGAGATGACTTC 1129
|||||

QY 1180 atccccctgcttccacagtgacggggccagggtgatgtgca 1223
|||||

Db 1130 ATCCCCCTGCTCCACAGTGACGGGGCCAGGGTGATGTGCA 1173
|||||

RESULT 9

I78503 98 bp DNA PAT 03-APR-1998
LOCUS
DEFINITION Sequence 14 from patent US 5693756.
ACCESSION I78503
VERSION I78503.1 GI:3014557
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 98)
AUTHORS Li,X., Blackshaw,S. and Snyder,S.H.
TITLE Amiloride-sensitive sodium channel and method of identifying substances which stimulate or block salty taste perception
JOURNAL Patent: US 5693756-A 14 02-DEC-1997;
FEATURES Location/Qualifiers
Source 1..98

BASE COUNT 18 a 25 c 26 g 29 t
ORIGIN

Query Match 3.1%; Score 98; DB 5; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.4e-46;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1517 cctgggctattgctattataaactcagggcgccctctccttgacagcctggctgtt 1576
|||||

Db 1 CCTGGGGCTATTGCTATTATAACTCGAGGGCGCCTCTCCTTGGACAGCTGGGCTGTT 60
|||||

QY 1577 tctccaagtgtggaagcctttagtgtgatcaactac 1614
|||||

Db 61 TCTCCAAGTGTGGAAGCCTTGTAGTGTGATCACTAC 98
|||||

RESULT 10

AF002665 852 bp DNA ROD 05-JAN-1999
LOCUS
DEFINITION Rattus norvegicus epithelial sodium channel alpha subunit (Scnnla)
ACCESSION AF002665
VERSION AF002665.1 GI:4100903
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 852)
AUTHORS Stock,P., Kreutz,R., Ganten,D. and Lindpaintner,K.
TITLE Identification of two microsatellites in the gene coding for the alpha epithelium sodium channel subunit in the rat
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 852)
AUTHORS Kreutz,R.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1997) Clinical Pharmacology, Benjamin Franklin Hospital, Free University of Berlin, Hindenburgdamm 30, Berlin D-12200, Germany

FEATURES

source 1..852
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Stroke-prone spontaneously hypertensive rat (SHRSPHD), Heidelberg strain"
/db_xref="taxon:10116"
/chromosome="4"
/map="4q42: Spr-D4Mit20-Eno2-Scnnla"
/clone="pCRII-8-Scnnla"
<1..>114
/gene="Scnnla"
/product="epithelial sodium channel alpha subunit"
<1..>114
/gene="Scnnla"
<1..>114
/note="RNASNAAC; RNEENACA; similar to epithelial sodium channel alpha subunits encoded by GenBank Accession Numbers X70497 and X70521"
/codon_start=1
/product="epithelial sodium channel alpha subunit"
/protein_id="AAD01004.1"
/db_xref="GI:4100904"
/translation="LRTPPPPYSGRTARGSSSVDRNNPQVDRKWKIGQL"

mRNA

gene

CDS

BASE COUNT 202 a 200 c 253 g 197 t
ORIGIN

Query Match 2.3%; Score 71; DB 12; Length 852;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 776 gcgggtcttcagctgcgcacaaatcccccaagtgagccggaagactggaagatcg 835
|||||

Db 44 GCGGGTCTTCAGGTACGCGACAAATCCCCAAGTGGACCGGAGGACTGGAAGATCG 103
|||||

QY 836 gcttccaactg 846
|||||

Db 104 GCTTCCAACATG 114
|||||

RESULT 11

CP0249296
LOCUS
DEFINITION Cavia porcellus mRNA for sodium channel, alpha subunit.
ACCESSION AJ249296
VERSION AJ249296.1 GI:6687302
KEYWORDS alpha subunit; sodium channel.
SOURCE domestic guinea pig.
ORGANISM Cavia porcellus

REFERENCE 1 (bases 1 to 2695)
AUTHORS Schnizler,M., Mastroberardino,L., Reifarth,F., Weber,W.M., Verrey,F. and Claus.W.

TITLE CAMP sensitivity conferred to the epithelial Na+ channel by alpha-subunit cloned from guinea-pig colon
JOURNAL Pflugers Arch. 439 (5), 579-587 (2000)
MEDLINE 2025288
REFERENCE 2 (bases 1 to 2695)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

AUTHORS Schnizler, M.K.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Schnizler M.K., Biologie, Institut fuer
Tierphysiologie, Justus-Liebig-Universitaet, Wartweg 95, 35392
Giessen, GERMANY

FEATURES
source 1. .2695
Location/Qualifiers
/organism="Cavia porcellus"
/db_xref="taxon:10141"
/country="Germany"
/sex="male"
/dev_stage="adult"
/tissue_type="colonic mucosa"
/tissue_lib="colonic mucosa"
293. .2263
/function="epithelial sodium channel"
/codon_start=1
/product="alpha subunit"
/protein_id="CAB64910.1"
/db_xref="GI:6687303"
/translation="MKGDELKEOGLPPPOLOGPLKDKCEQPLGPEPTAPOHTEE
EEALTEFHSRYELFOFFCNNTIHGAILRVCSKHNRMKTAFAWVLMCTFGMYWQF
ALLFGYFSPVSNLNLNSDKLVFAVTVCTLNRYKEIKQELRELDRIQTQTLFD
LYNNASSTLLAGARRSLADTLPLQPIVQPEPRARRSDPSVRDNNPRVDR
DMRVGQLCNQKSDCFYQTSVGVDGVRWFHYINILAOVADTSPSLEELALGNF
IFACRNOACTOENTSHPHPIYGNCTYFNKNDSLLWMSMPGINSLSLTLRTEQ
NDYIPLLSVTGARVTVHGODEFAFMDDGFGNLRPGVETISMRKEALDRLGSGYDC
TQGSVPQVNLPSKYTQVCIHSCFOENMKQCCAYIFPKPGVQPCDTRKQSA
WGCTYKLGAGSSDLSGCKNCKPCNVITYIKLSAGYSRWPSTVSDWTFQWLSLN
NYTISKRGVANLTYFELNYRTNSESPTVMVLLSLNLSQWLSWFGSSVLSVYE
MAFMFDLLVITLLMLRFRSRYWSPGRGARAARVACTPPPSLPSRFCAHSAFPTL
TAPPPAYATLSACPPLOGLAGASSACAPREP"

BASE COUNT 577 a 855 c 754 g 509 t
ORIGIN

Query Match 1.9%; Score 59; DB 12; Length 2695;
Best Local Similarity 100.0%; Pred. No. 4.2e-23;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1786 gtacagatggtcagctctgtccaaactggcagcagcagtgagcctgtgttggctc 1844
|||||
Db 1910 GTCACGATGTCAGCCCTCTCCAACTGGCGAGCCAGTCAGGCTGTGTTGGCTC 1968
|||||

RESULT 12
I78500/c
LOCUS 178500 45 bp DNA PAT 03-APR-1998
DEFINITION Sequence 11 from patent US 5693756.
ACCESSION I78500
VERSION I78500.1 GI:3014654
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 45)
AUTHORS Li, X., Blackshaw, S. and Snyder, S.H.
TITLE Amiloride-sensitive sodium channel and method of identifying
substances which stimulate or block salty taste perception
JOURNAL Patent: US 5693756-A 11 02-DEC-1997;
FEATURES Location/Qualifiers
source 1. .45
/organism="unknown"
BASE COUNT 10 a 10 c 12 g 13 t
ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 5.4e-15;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1134 caacaatggtttccctgacactgcgcacagacagaatgactt 1178
|||||

Db 45 CAACAATGTTTGTCCCTGACACTGCGCACAGACAGAAATGACTT 1

RESULT 13
I78501/c
LOCUS 178501 45 bp DNA PAT 03-APR-1998
DEFINITION Sequence 12 from patent US 5693756.
ACCESSION I78501
VERSION I78501.1 GI:3014655
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 45)
AUTHORS Li, X., Blackshaw, S. and Snyder, S.H.
TITLE Amiloride-sensitive sodium channel and method of identifying
substances which stimulate or block salty taste perception
JOURNAL Patent: US 5693756-A 12 02-DEC-1997;
FEATURES Location/Qualifiers
source 1. .45
/organism="unknown"
BASE COUNT 11 a 11 c 10 g 13 t
ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 5.4e-15;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1498 gactaccgaagcagagctcctgggctattgtctattataaactg 1542
|||||

Db 45 GACTACCGAAAGCAGAGCTCTCGGGCTATTGCTATTATAAATG 1

RESULT 14
OCU132108
LOCUS OCU132108 2915 bp mRNA MAM 13-JAN-1999
DEFINITION Oryctolagus cuniculus mRNA for epithelial sodium channel, alpha
subunit.
ACCESSION AJ132108
VERSION AJ132108.1 GI:4158223
KEYWORDS alpha subunit; alpha-ENaC gene; epithelial sodium channel.
SOURCE Oryctolagus cuniculus.
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 2915)
AUTHORS Kudlacek, O.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1999) Kudlacek O., Pharmacological Institute,
University of Vienna, Waehringerstrasse 13A, A-1090 Wien, AUSTRIA
REFERENCE 2 (bases 1 to 2915)
AUTHORS Kudlacek, O., Weisz, E., Wiener, H. and Plass, H.
TITLE The rabbit epithelial sodium channel
JOURNAL Unpublished
FEATURES Location/Qualifiers
source 1. .2915
/organism="Oryctolagus cuniculus"
/db_xref="taxon:9986"
35. .1957
/gene="alpha-ENaC"
35. .1957
/gene="alpha-ENaC"
/codon_start=1
/product="epithelial sodium channel, alpha subunit"
/protein_id="CAA10571.1"
/db_xref="GI:4158224"
/translation="MKGDKREOGCPGPEVAPQPTAEALIEFHSRYELFOFFCN
NTIHGAIRLVCSKHNRMKTAFAWVLMCTFGMYWQFGLLFGFYFSPYNLNLS
DKLVFPVAVTVCTLNRYPEITDOLKELSDITQTLDLFLKYNASTLEAQRHRRDVH
PLPLPHQLRVLPPPPRLEARRASSSDNSPEVGRKDMWIGFQLCNQNRSDCFYQ
RYSGGVDVAREVRFHYINILSLSDTSLSRQLGNFTTCRFNQAFCGDGNYSHFHH
PMYGCNYTFNDKNNSLWMSMPGINSLSLTLRTEQNDFTPLLSVTGTGARVMVHGQD

EPAPMDGGFNLRGCVETISMRKESLDRLGGDYDCTQNGSDVPVKNLYRSKYTQOV
CIHSFENVMYKEGCAYIFYPLPEGVYCDYRKHNSWCYCYKLODAFSSDRLGCFE
KCRKPCSVNTYELSGYSRPSVTSDQVYFOMLSIQNNYTVSNKRNGVAKNIYFKEL
NYKANSSEPSVTMYTLISNLGSSQSLAFSGSVLSVMEAEILFDLSVITFMLLRFR
SRWSPGRGAGGAREVASSPVSAFPSRCPHTSPSPQPGFTLPPSLATPPAYAIL
GPCUSQSGSACAPGP"
BASE COUNT 577 a 978 c 774 g 586 t
ORIGIN

Query Match 1.3%; Score 41; DB 3; Length 2915;
Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 325 gagctctccagttcttgcacaacaccaccatccacgg 365
|||||
Db 143 GAGCTCTCCAGTCTCTTCGACACACACCACCATCCACGG 183
|||||

RESULT 15
PAU95179 636 bp mRNA PLN 15-DEC-1997
LOCUS Prunus armeniaca putative Nifu protein mRNA, partial cds.
DEFINITION
ACCESSION U95179
VERSION U95179.1 GI:2688825
KEYWORDS
SOURCE apricot.
ORGANISM

Prunus armeniaca
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core
eudicots; Rosidae; eurosids I; Rosales; Rosaceae; Prunus.
REFERENCE 1 (bases 1 to 636)
AUTHORS Mbequie A Mbequie,D., Gomez,R.-M. and Fills-Lycaon,B.
TITLE Molecular cloning and nucleotide sequence of a putative Nifu
protein from apricot
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 636)
AUTHORS Mbequie A Mbequie,D., Gomez,R.-M. and Fills-Lycaon,B.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-1997) Station de Technologie des Produits
Vegetaux, INRA, Site AGRORPARC, domaine Saint Paul, Avignon 84914
cedex 9, France

FEATURES
source
1. .636
/organism="Prunus armeniaca"
/strain="Bergeron"
/db_xref="taxon:36596"
/tissue_type="mesocarp; exocarp"
/dev_stage="ripe fruit"
/clone="pAPRI31"
<1. .231
/codon_start=1
/product="putative Nifu protein"
/protein_id="AAB88877.1"
/db_xref="GI:2688826"
/translation="ELAVPNVDVLEDRPYLIADGGVDVSVSDGVVSLKQGCAGC
SCPSTTTMKGIERVLKEKFGDALKDIQQL"
BASE COUNT 199 a 118 c 149 g 170 t
ORIGIN

Query Match 1.3%; Score 40; DB 8; Length 636;
Best Local Similarity 100.0%; Pred. No. 6.4e-12;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 581 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 620
|||||

RESULT 16
GGVITRO 1620 bp mRNA VRT 04-DEC-1997
LOCUS

G.gallus mRNA for vitronectin.
ACCESSION Y11030
VERSION Y11030.1 GI:1922281
KEYWORDS vitronectin.
SOURCE chicken.
ORGANISM

Gallus gallus
Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1620)
AUTHORS Martinez-Morales,J.R., Barbas,J.A., Marti,E., Bovolenta,P.,
Egarr,D. and Rodriguez-rebar,A.

TITLE Vitronectin is expressed in the ventral region of the neural tube
and promotes the differentiation of motor neurons
JOURNAL Development 124 (24), 5139-5147 (1997)
MEDLINE 98090066
REFERENCE 2 (bases 1 to 1620)
AUTHORS Barbas,J.A.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-1997) J.A. Barbas, Instituto Cajal CSIC, Av.
Doctor Acre 37, E- 28002 Madrid, SPAIN

FEATURES
source
1. .1620
/organism="Gallus gallus"
/strain="White leghorn"
/db_xref="taxon:9031"
/dev_stage="7 day embryo"
/tissue_type="retina"
/clone_lib="lambda-ZAP II"
141. .1502
/codon_start=1
/product="vitronectin"
/protein_id="CAA71914.1"
/db_xref="GI:1922282"
/translation="MRLLPVLVLLALTRAEDSCGRDEGNAMKKKCCDILCN
YYQCCSDYSTVCRAKVTGDFVLPEDDYLDYLSIDTGVGTEAPEAPTEHPTEPH
SSPPTRVIDTATEETPEPEVPTLHPTTTTTSDTRNPSLDEPEELCSRKPFN
AFTDLKNGSIYAFKGYELDKSVRGYKLIISDVWIGIEGPIDAATFRVNCQKTY
LFKGSYWRFDGALDPCYPRDISEGFIGIPNDIDAAFPALPAHSYGHVGRVYFKGKY
YWSDFAHQPTQAECEKSSPSTVFNHYAFMRDSDHEDIFLSIFGSRWVGASORLISR
DWGVPNLDAAAGRIYVSSRQRRSRHRKRNRHTLNGLWLSWLNDSSESTD
TESDWLSGSCQETLQSVYFFVGDKYRYVNLRTKRVLDVOPPYPRISIAQYWLDCQPQDE
EST"
polyA_signal 1550. .1555
BASE COUNT 399 a 510 c 431 g 280 t
ORIGIN

Query Match 1.3%; Score 40; DB 4; Length 1620;
Best Local Similarity 100.0%; Pred. No. 7e-12;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 1570 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1609
|||||

RESULT 17
AB002533 2171 bp mRNA PRI 13-FEB-1999
LOCUS Homo sapiens mRNA for Qip1, complete cds.
DEFINITION
ACCESSION AB002533
VERSION AB002533.1 GI:1944124
KEYWORDS Qip1.
SOURCE Homo sapiens cell_line:HeLa S3 cDNA to mRNA.
ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2171)
AUTHORS Seki,T.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-1997) to the DDBJ/EMBL/GenBank databases.
Takahiko Seki, Tohoku University, Faculty of Pharmaceutical

Sciences; Aoba Aramaki, Aoba-Ku, Sendai, Miyagi 980-77, Japan
(E-mail: takaeghi2.pharm.tohoku.ac.jp, Tel: 81-22-217-6876,
Fax: 81-22-217-6873)
2 (sites)
Seki, T., Tada, S., Katada, T. and Enomoto, T.
Cloning of a cDNA encoding a novel importin-alpha homologue, Qip1:
discrimination of Qip1 and Rch1 from hsrp1 by their ability to
interact with DNA helixase Q1/RecQ1
Biochem. Biophys. Res. Commun. (1997) In press

-----Summary Statistics
Consensus quality: 47661 bases at least Q40
Consensus quality: 55452 bases at least Q30
Consensus quality: 57169 bases at least Q20
Estimated insert size: 61001; sum-of-contigs estimation
Estimated insert size: 61001; pulse field gel estimation
Quality coverage: 3.20x in Q20 bases; pulse field gel estimation
Quality coverage: 3.46x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1624: contig of 1624 bp in length
* gap of unknown length
* 1625 2807: contig of 1183 bp in length
* gap of unknown length
* 2808 4042: contig of 1235 bp in length
* gap of unknown length
* 4043 5495: contig of 1453 bp in length
* gap of unknown length
* 5496 7142: contig of 1647 bp in length
* gap of unknown length
* 7143 9094: contig of 1952 bp in length
* gap of unknown length
* 9095 10968: contig of 1874 bp in length
* gap of unknown length
* 10969 13357: contig of 2389 bp in length
* gap of unknown length
* 13358 15997: contig of 2640 bp in length
* gap of unknown length
* 15998 20030: contig of 4033 bp in length
* gap of unknown length
* 20031 24641: contig of 4611 bp in length
* gap of unknown length
* 24642 28743: contig of 4102 bp in length
* gap of unknown length
* 28744 33674: contig of 4931 bp in length
* gap of unknown length
* 33675 41268: contig of 7594 bp in length
* gap of unknown length
* 41269 51931: contig of 10863 bp in length
* gap of unknown length
* 51932 61001: contig of 9070 bp in length.
* Location/Qualifiers
1. 61001
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTB-192J21"
BASE COUNT 16555 a 14407 c 13948 g 16090 t 1 others
ORIGIN

Query Match 1.3%; Score 40; DB 9; Length 2171;
Best Local Similarity 100.0%; Pred. No. 7.2e-12;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 1678 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1717

RESULT 18
AC034204
LOCUS AC034204 61001 bp DNA HTG 09-APR-2000
DEFINITION Homo sapiens chromosome 5 clone CTB-192J21, WORKING DRAFT SEQUENCE,
16 unordered pieces.
AC034204
VERSION AC034204.2 GI:7528346
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 61001)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
2 (bases 1 to 61001)
REFERENCE DOE Joint Genome Institute.
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Apr 9, 2000 this sequence version replaced gi:7417681.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

polyA_site
/note="32 A nucleotides"
BASE COUNT 685 a 450 c 477 g 559 t
ORIGIN

Query Match 1.3%; Score 40; DB 9; Length 2171;
Best Local Similarity 100.0%; Pred. No. 7.2e-12;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 1678 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1717

RESULT 18
AC034204
LOCUS AC034204 61001 bp DNA HTG 09-APR-2000
DEFINITION Homo sapiens chromosome 5 clone CTB-192J21, WORKING DRAFT SEQUENCE,
16 unordered pieces.
AC034204
VERSION AC034204.2 GI:7528346
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 61001)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
2 (bases 1 to 61001)
REFERENCE DOE Joint Genome Institute.
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Apr 9, 2000 this sequence version replaced gi:7417681.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

VERSION KEYWORDS SOURCE

AC036192.1 GI:7523861
HTG: HTGS_PHASE0.

human.

ORGANISM

Homo sapiens

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.

REFERENCE

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens chromosome 2, clone RP11-485P2

JOURNAL

REFERENCE

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye.W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (07-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L9272

Center clone name: 485_P_2

* NOTE: This record contains 89 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 804: contig of 804 bp in length
* 805 904: gap of 100 bp
* 905 1712: contig of 808 bp in length
* 1713 1812: gap of 100 bp
* 1813 2550: contig of 738 bp in length
* 2551 2650: gap of 100 bp
* 2651 3426: contig of 776 bp in length
* 3427 3526: gap of 100 bp
* 3527 4336: contig of 810 bp in length
* 4337 4436: gap of 100 bp
* 4437 5244: contig of 808 bp in length
* 5245 5344: gap of 100 bp
* 5345 6160: contig of 816 bp in length
* 6161 6260: gap of 100 bp
* 6261 7064: contig of 804 bp in length

* 7065 7164: gap of 100 bp
* 7165 7973: contig of 809 bp in length
* 7974 8073: gap of 100 bp
* 8074 8896: contig of 823 bp in length
* 8897 8996: gap of 100 bp
* 8997 9801: contig of 805 bp in length
* 9802 9901: gap of 100 bp
* 9902 10730: contig of 829 bp in length
* 10731 10830: gap of 100 bp
* 10831 11625: contig of 795 bp in length
* 11626 11725: gap of 100 bp
* 11726 12520: contig of 795 bp in length
* 12521 12620: gap of 100 bp
* 12621 13419: contig of 799 bp in length
* 13420 13519: gap of 100 bp
* 13520 14328: contig of 809 bp in length
* 14329 14428: gap of 100 bp
* 14429 15246: contig of 818 bp in length
* 15247 15346: gap of 100 bp
* 15347 16151: contig of 805 bp in length
* 16152 16251: gap of 100 bp
* 16252 17071: contig of 820 bp in length
* 17072 17171: gap of 100 bp
* 17172 17969: contig of 798 bp in length
* 17970 18069: gap of 100 bp
* 18070 18859: contig of 790 bp in length
* 18860 18959: gap of 100 bp
* 18960 19786: contig of 827 bp in length
* 19787 19886: gap of 100 bp
* 19887 20724: contig of 838 bp in length
* 20725 20824: gap of 100 bp
* 20825 21607: contig of 783 bp in length
* 21608 21707: gap of 100 bp
* 21708 22533: contig of 826 bp in length
* 22534 22633: gap of 100 bp
* 22634 23446: contig of 813 bp in length
* 23447 23546: gap of 100 bp
* 23547 24331: contig of 785 bp in length
* 24332 24431: gap of 100 bp
* 24432 25256: contig of 825 bp in length
* 25257 25356: gap of 100 bp
* 25357 26163: contig of 807 bp in length
* 26164 26263: gap of 100 bp
* 26264 27095: contig of 832 bp in length
* 27096 27195: gap of 100 bp
* 27196 28020: contig of 825 bp in length
* 28021 28120: gap of 100 bp
* 28121 28945: contig of 825 bp in length
* 28946 29045: gap of 100 bp
* 29046 29878: contig of 833 bp in length
* 29879 29978: gap of 100 bp
* 29979 30778: contig of 800 bp in length
* 30779 30878: gap of 100 bp
* 30879 31648: contig of 770 bp in length
* 31649 31748: gap of 100 bp
* 31749 32549: contig of 801 bp in length
* 32550 32649: gap of 100 bp
* 32650 33472: contig of 823 bp in length
* 33473 33572: gap of 100 bp
* 33573 34389: contig of 817 bp in length
* 34390 34489: gap of 100 bp
* 34490 35310: contig of 821 bp in length
* 35311 35410: gap of 100 bp
* 35411 36211: contig of 801 bp in length
* 36212 36311: gap of 100 bp
* 36312 37139: contig of 828 bp in length
* 37140 37239: gap of 100 bp
* 37240 38052: contig of 813 bp in length
* 38053 38152: gap of 100 bp
* 38153 38986: contig of 834 bp in length
* 38987 39086: gap of 100 bp
* 39087 39924: contig of 838 bp in length
* 39925 40024: gap of 100 bp

* 40025 40808: contig of 784 bp in length
* 40809 40908: gap of 100 bp
* 40909 41725: contig of 817 bp in length
* 41726 41825: gap of 100 bp
* 41826 42620: contig of 795 bp in length
* 42621 42720: gap of 100 bp
* 42721 43526: contig of 806 bp in length
* 43527 43626: gap of 100 bp
* 43627 44435: contig of 809 bp in length
* 44436 44535: gap of 100 bp
* 44536 45349: contig of 814 bp in length
* 45350 45449: gap of 100 bp
* 45450 46248: contig of 799 bp in length
* 46249 46348: gap of 100 bp
* 46349 47147: contig of 799 bp in length
* 47148 47247: gap of 100 bp
* 47248 48063: contig of 816 bp in length
* 48064 48163: gap of 100 bp
* 48164 48984: contig of 821 bp in length
* 48985 49084: gap of 100 bp
* 49085 49891: contig of 807 bp in length
* 49892 49991: gap of 100 bp
* 49992 50801: contig of 810 bp in length
* 50802 50901: gap of 100 bp
* 50902 51691: contig of 790 bp in length
* 51692 51791: gap of 100 bp
* 51792 52610: contig of 819 bp in length
* 52611 52710: gap of 100 bp
* 52711 53528: contig of 818 bp in length
* 53529 53628: gap of 100 bp
* 53629 54440: contig of 812 bp in length
* 54441 54540: gap of 100 bp
* 54541 55346: contig of 806 bp in length
* 55347 55446: gap of 100 bp
* 55447 56243: contig of 797 bp in length
* 56244 56343: gap of 100 bp
* 56344 57148: contig of 805 bp in length
* 57149 57248: gap of 100 bp
* 57249 58063: contig of 815 bp in length
* 58064 58163: gap of 100 bp
* 58164 58983: contig of 820 bp in length
* 58984 59083: gap of 100 bp
* 59084 59899: contig of 816 bp in length
* 59900 59999: gap of 100 bp
* 60000 60812: contig of 813 bp in length
* 60813 60912: gap of 100 bp
* 60913 61734: contig of 822 bp in length
* 61735 61834: gap of 100 bp
* 61835 62644: contig of 810 bp in length
* 62645 62744: gap of 100 bp

Query Match 1.3%; Score 40; DB 73; Length 80861;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tcctatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 54838 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 54877

RESULT 20
AC026450/c
LOCUS AC026450 95236 bp DNA HTG 22-MAR-2000
DEFINITION Homo sapiens chromosome 5 clone CTD-2381N15, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
ACCESSION AC026450
VERSION AC026450.1 GI:7279615
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 95236)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
2 (bases 1 to 95236)
DOE Joint Genome Institute.
Direct Submission
Submitted (22-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>
-----Summary Statistics
Consensus quality: 83275 bases at least Q40
Consensus quality: 89876 bases at least Q30
Consensus quality: 91237 bases at least Q20
Estimated insert size: 95236; sum-of-contigs estimation
Estimated insert size: 83310; agarose-fp estimation
Quality coverage: 4.77x in Q20 bases; agarose-fp estimation
Quality coverage: 4.17x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1090: contig of 1090 bp in length
* gap of unknown length
* 1091 2567: contig of 1477 bp in length
* gap of unknown length
* 2568 3633: contig of 1066 bp in length
* gap of unknown length
* 3634 5602: contig of 1969 bp in length
* gap of unknown length
* 5603 7519: contig of 1917 bp in length
* gap of unknown length
* 7520 11811: contig of 4292 bp in length
* gap of unknown length
* 11812 17036: contig of 5225 bp in length
* gap of unknown length
* 17037 21227: contig of 4191 bp in length
* gap of unknown length
* 21228 27367: contig of 6140 bp in length
* gap of unknown length
* 27368 35548: contig of 8181 bp in length
* gap of unknown length
* 35549 44349: contig of 8801 bp in length
* gap of unknown length
* 44350 52374: contig of 8025 bp in length
* gap of unknown length
* 52375 63936: contig of 11562 bp in length
* gap of unknown length
* 63937 81454: contig of 17518 bp in length
* gap of unknown length
* 81455 95236: contig of 13782 bp in length.

FEATURES
source

1. .95236
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2381N15"
BASE COUNT 27883 a 19363 c 19396 g 28584 t 10 others
ORIGIN

Query Match 1.3%; Score 40; DB 67; Length 95236;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tcctatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

[illegible]

RESULT 21

HSDJ900E8/C

LOCUS

DEFINITION

Contains the first coding exon of the GPC4 gene for glypican 4, STSS, GSSs and a putative CpG island, complete sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCES

REFERENCE

REFERENCE AUTHORS

TITLE

```

#####
CCATCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2013

HSDJ900E8 101259 bp DNA PRI 07-APR-2000
Human DNA sequence from clone RP5-900E8 on chromosome Xq25-27.1.
Contains the first coding exon of the GPC4 gene for glypican 4,
STSS, cSSs and a putative CpG island, complete sequence.
AL109623
AL109623.9 GI:7529584
HTG: CpG island; glypican; GPC4.
human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 101259)
Wilson.S.
Direct Submission
Submitted (21-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
Requests: clonerequest@sanger.ac.uk
On Apr 9, 2000 this sequence version replaced gi:6015559.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP5-900E8 is from the library RPc1-5 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCIPAC2

IMPORTANT: This sequence is not the entire insert of clone
RP5-900E8 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP5-900E8 is at 101259 in this
sequence. The true right end of clone RP6-198C21 is at 100 in this
sequence.

Location/Qualifiers
1..101259
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q25-27.1"
/clone="RP5-900E8"
/clone_lib="RPc1-5"
1..318
/region
/note="MULTIAl repeat: matches 21..359 of consensus"
395..449
/region
/note="L2 repeat: matches 2634..2689 of consensus"
487..780
/region
/note="AluX repeat: matches 1..295 of consensus"
1050..1352
/region
/note="AluSg repeat: matches 1..302 of consensus"
1363..1652
/region
/note="AluSg repeat: matches 1..299 of consensus"

```

```
repeat_region /note="AluSg repeat: matches 1. .139 of consensus"
22822. .23144
repeat_region /note="AluJb repeat: matches 1. .308 of consensus"
23557. 23828
repeat_region /note="AluSg1 repeat: matches 5. .301 of consensus"
24093. .24399
repeat_region /note="AluJb repeat: matches 1. .307 of consensus"
24403. .24662
repeat_region /note="MIR repeat: matches 12. .262 of consensus"
24690. 24770
repeat_region /note="L1R15 repeat: matches 172. .248 of consensus"
24790. .25093
repeat_region /note="AluSg repeat: matches 1. .303 of consensus"
25129. .25628
repeat_region /note="L1M85 repeat: matches 5646. .6143 of consensus"
25637. 25938
repeat_region /note="AluSx repeat: matches 5. .306 of consensus"
25967. .26249
repeat_region /note="AluJb repeat: matches 1. .282 of consensus"
26279. .26567
repeat_region /note="AluJb repeat: matches 7. .281 of consensus"
Complement(26853. .27267)
misc_feature /note="match: GSS: Em:AQ430698"
Complement(26889. .27291)
misc_feature /note="match: GSS: Em:AQ496873"
28401. .28584
repeat_region /note="MER30 repeat: matches 1. .197 of consensus"
29736. .30259
misc_feature /note="match: GSS: Em:B99950"
29945. .30071
repeat_region /note="MIR repeat: matches 34. .165 of consensus"
30073. .30336
repeat_region /note="AluSg repeat: matches 1. .281 of consensus"
30883. .30941
repeat_region /note="AluJb repeat: matches 86. .139 of consensus"
30942. .31601
repeat_region /note="MER51B repeat: matches 1. .617 of consensus"
31602. .31737
repeat_region /note="AluJb repeat: matches 139. .297 of consensus"
31877. .32185
repeat_region /note="AluJb repeat: matches 1. .299 of consensus"
32660. .32766
repeat_region /note="L2 repeat: matches 2634. .2744 of consensus"
33288. .33581
repeat_region /note="AluSx repeat: matches 4. .296 of consensus"
34977. .35285
repeat_region /note="AluJb repeat: matches 7. .310 of consensus"
35360. .35429
repeat_region /note="MIR repeat: matches 93. .162 of consensus"
35750. .35837
repeat_region /note="44 copies 2 mer tt 65% conserved"
35863. .36172
repeat_region /note="AluJb repeat: matches 1. .312 of consensus"
36206. .36502
repeat_region /note="AluSc repeat: matches 1. .308 of consensus"
36921. .37101
repeat_region /note="MIR repeat: matches 76. .255 of consensus"
38901. .39098
repeat_region /note="MER3 repeat: matches 1. .204 of consensus"
39140. .39430
```

```
Query Match 1.3%; Score 40; DB 11; Length 101259;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 58491 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 58452
```

```
RESULT 22
AC020928
LOCUS AC020928 103804 bp DNA HTG 06-APR-2000
```

```
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

```
Homo sapiens chromosome 19 clone CTD-2162K18, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
AC020928
AC020928.3 GI:7458788
HTG: HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 103804)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 19
Unpublished
2 (bases 1 to 103804)
DOE Joint Genome Institute.
Direct Submission
Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 6, 2000 this sequence version replaced gi:7021678.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----Summary Statistics
Consensus quality: 96332 bases at least Q40
Consensus quality: 101562 bases at least Q30
Consensus quality: 102259 bases at least Q20
Estimated insert size: 103804; sum-of-contigs estimation
Estimated insert size: 111350; agarose-fp estimation
Quality coverage: 4.80x in Q20 bases; agarose-fp estimation
Quality coverage: 5.15x in Q20 bases; sum-of-contigs estimation
-----
```

* NOTE: This is a 'working draft' sequence. It currently

* consists of 10 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved. 2062: contig of 2062 bp in length

* 1

* 2063 6764: contig of 4702 bp in length

* gap of unknown length

* 6765 12879: contig of 6115 bp in length

* gap of unknown length

* 12880 19068: contig of 6189 bp in length

* gap of unknown length

* 19069 23463: contig of 4395 bp in length

* gap of unknown length

* 23464 35319: contig of 11856 bp in length

* gap of unknown length

* 35320 47360: contig of 12041 bp in length

* gap of unknown length

* 47361 61652: contig of 14292 bp in length

* gap of unknown length

* 61653 81957: contig of 20305 bp in length

* gap of unknown length

* 81958 103804: contig of 21847 bp in length.

Location/Qualifiers

1. .103804

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="19"

/clone="CTD-2162K18"

FEATURES

source

BASE COUNT 30040 a 22650 c 22148 g 28963 t 3 others

ORIGIN

Query Match 1.3%; Score 40; DB 73; Length 103804;

Best Local Similarity 100.0%; Pred. No. 1.1e-11;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 92154 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 92193

RESULT 23

AC007880 AC007880 122223 bp DNA PRI 21-DEC-1999
DEFINITION Homo sapiens BAC clone RP11-323F11 from 2, complete sequence.
AC007880
VERSION AC007880.2 GI:5931460
KEYWORDS HTG.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792

REFERENCE

AUTHORS Du, H., Maupin, R. A. C. H. E. L., Yeakum, M. and Nguyen, C.
TITLE The sequence of Homo sapiens BAC clone RP11-323F11
JOURNAL Unpublished

REFERENCE

AUTHORS Waterston, R. H.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

AUTHORS Waterston, R. H.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE

AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 28, 1999 this sequence version replaced gi:5103899.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0323F11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis

MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-359K10, 200 bp overlap; the clone sequenced to the right is RP11-451M22. Actual start of this clone is at base position 147528 of RP11-359K10.

FEATURES

	Location/Qualifiers
source	1. .122223
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-323F11"
	/clone_lib="RPCI-11"
repeat_region	694. .804
	/rpt_family="MER1_type"
repeat_region	1020. .1322
	/rpt_family="Alu"
repeat_region	1367. .1661
	/rpt_family="Alu"
repeat_region	1792. .1960
	/rpt_family="Retroviral"
repeat_region	1957. .2161
	/rpt_family="Retroviral"
repeat_region	4727. .4751
	/rpt_family="(TTA)n"
repeat_region	4752. .10803
	/rpt_family="L1"
repeat_region	13379. .13905
	/rpt_family="Retroviral"
repeat_region	13902. .13962
	/rpt_family="Harlequin"
repeat_region	13959. .14032
	/rpt_family="Harlequin"
repeat_region	14033. .14196
	/rpt_family="Retroviral"
repeat_region	14200. .14318
	/rpt_family="Retroviral"
repeat_region	14305. .14415
	/rpt_family="Retroviral"
repeat_region	15049. .15114
	/rpt_family="MER1_type"
misc_feature	15105. .15485
	/note="match to EST A1809723 (NID:G5396289) wh77a01.x1"
repeat_region	15132. .15244
	/rpt_family="MIR"
repeat_region	15447. .15775
	/rpt_family="L1"
repeat_region	15806. .16161
	/rpt_family="Retroviral"
repeat_region	16162. .16361
	/rpt_family="L1"
repeat_region	16362. .16526
	/rpt_family="MaLR"
repeat_region	16527. .16669
	/rpt_family="Alu"
repeat_region	16670. .16981
	/rpt_family="Alu"
repeat_region	16982. .17147
	/rpt_family="Alu"
repeat_region	17148. .17317
	/rpt_family="MaLR"

```
repeat_region 17318..17623
/rpt_family="Alu"
repeat_region 17624..17693
/rpt_family="MaIR"
repeat_region 17694..17806
/rpt_family="LI"
repeat_region 17807..17829
/rpt_family="LI"
repeat_region 17830..18105
/rpt_family="(TTAA)n"
repeat_region 18106..18159
/rpt_family="Alu"
repeat_region 18227..18525
/rpt_family="LI"
misc_feature 19214..19349
/rpt_family="Alu"
/note="match to EST T81027 (NID:g703912) yd25c02.sl"
repeat_region 19310..19420
/rpt_family="LI"
repeat_region 19421..19704
/rpt_family="Alu"
repeat_region 19705..19747
/rpt_family="LI"
repeat_region 19770..20037
/rpt_family="LI"
repeat_region 20178..20286
/rpt_family="L2"
repeat_region 20304..20449
/rpt_family="MER2_type"
repeat_region 20721..21260
/rpt_family="Retroviral"
repeat_region 21790..21899
/rpt_family="MIR"
repeat_region 22010..22404
/rpt_family="L2"
repeat_region 23178..23398
/rpt_family="LI"
repeat_region 23399..23826
/rpt_family="LI"
repeat_region 23827..24134
/rpt_family="LTR54"
repeat_region 24135..24243
/rpt_family="Alu"
repeat_region 24244..24471
/rpt_family="LTR54"
repeat_region 24472..24764
/rpt_family="LI"
repeat_region 24765..24888
/rpt_family="Alu"
repeat_region 24907..25029
/rpt_family="LI"
repeat_region 25711..25787
/rpt_family="Alu"
repeat_region 26651..26675
/rpt_family="MER1_type"
repeat_region 26866..26926
/rpt_family="AT_rich"
repeat_region 26944..26985
/rpt_family="Alu"
repeat_region 26985..27031
/rpt_family="(GAAA)n"
repeat_region 27824..27945
/rpt_family="(GGGA)n"
```

```
Query Match 1.3%; Score 40; DB 39; Length 122223;
Best Local Similarity 100.0%; Pred. No. 1.1e-11; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0;
```

```
QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 70964 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 71003
```

```
RESULT 24
AC008932/c
```

```
LOCUS AC008932 128758 bp DNA HTG 05-APR-2000
DEFINITION Homo sapiens chromosome 5 clone CTD-2296H2, WORKING DRAFT SEQUENCE,
7 unordered pieces.
ACCESSION AC008932
VERSION AC008932.3 GI:7417522
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 128758)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 128758)
TITLE DOE Joint Genome Institute.
JOURNAL Direct Submission
COMMENT Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 5, 2000 this sequence version replaced gi:7025715.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----Summary Statistics
Consensus quality: 122924 bases at least Q40
Consensus quality: 126992 bases at least Q30
Consensus quality: 127709 bases at least Q20
Estimated insert size: 128758; sum-of-contigs estimation
Estimated insert size: 131000; pulse field gel estimation
Quality coverage: 5.00x in Q20 bases; pulse field gel estimation
Quality coverage: 5.09x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4080: contig of 4080 bp in length
* gap of unknown length
* 4081 9811: contig of 5731 bp in length
* gap of unknown length
* 9812 17645: contig of 7834 bp in length
* gap of unknown length
* 17646 26939: contig of 9294 bp in length
* gap of unknown length
* 26940 47546: contig of 20607 bp in length
* gap of unknown length
* 47547 69345: contig of 21799 bp in length
* gap of unknown length
* 69346 128758: contig of 59413 bp in length.
* Location/Qualifiers
1..128758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2296H2"
BASE COUNT 41820 a 22954 c 23096 g 40881 t 7 others
ORIGIN
```

FEATURES

Source

```
Query Match 1.3%; Score 40; DB 72; Length 128758;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 71426 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 71387
```


repeat_region 29218..29347
/note="L2 repeat: matches 2574..2703 of consensus"
repeat_region 29474..29546
/note="MSTD repeat: matches 327..394 of consensus"
repeat_region 29547..29859
/note="AluJb repeat: matches 1..306 of consensus"
repeat_region 29860..30156
/note="MSTD repeat: matches 1..327 of consensus"
repeat_region 33163..33464
/note="AluSx repeat: matches 1..302 of consensus"
repeat_region 34349..34666
/note="AluSx repeat: matches 1..312 of consensus"
misc_feature complement(<34667..>34935)
/note="match: GSS B43565"
misc_feature complement(<34823..35148)
/note="match: GSS AQ169413"
repeat_region 35018..35130
/note="MIR repeat: matches 33..143 of consensus"
repeat_region 35216..35521
/note="AluSg1 repeat: matches 1..303 of consensus"
repeat_region 36157..36204
/note="12 copies 4 mer acac 94% conserved"
repeat_region 36157..36202
/note="23 copies 2 mer ac 96% conserved"
misc_feature 36453..36901
/note="match: GSS AQ265693 clone 2510C6"
misc_feature 36453..36748
/note="match: GSS AQ265745 clone 2510C22"
misc_feature 36453..36751
/note="match: GSS AQ263897 clone 2510D5"
repeat_region 36765..37053
/note="AluJo repeat: matches 1..289 of consensus"
repeat_region 37520..37831
/note="AluSx repeat: matches 5..312 of consensus"
prim_transcript 38233..38731
/note="match: 3' EST AI417996 clone IMAGE:2114220"
repeat_region 38378..38401
/note="12 copies 2 mer ca 96% conserved"
repeat_region 38429..38481
/note="MADE1 repeat: matches 1..53 of consensus"
prim_transcript 38998..39479
/note="match: 3' EST AI368019 clone IMAGE:1934935"
prim_transcript 38998..39448
/note="match: 3' EST AI276057 clone IMAGE:1877575"
mRNA complement(join(<39482..41515,41612..41787,43171..43307,44218..44364,44470..44600,49704..49869,62591..62982,77629..>77787))
/note="match: AF030186; match: ESTs AA046087 AA046130 AA478468 AA893150 AA933954 AI071251 AI131068 AI141502 AI208484 AI385680 C82918 C83774 D19430 D25838 W12478 AA451958 C82589 C83445 C82710 C83566 W79980"
/evidence=not_experimental
repeat_region 40979..41026
/note="12 copies 4 mer acaa 75% conserved"
repeat_region 40979..41046
/note="34 copies 2 mer aa 66% conserved"

Query Match 1.3%; Score 40; DB 10; Length 141762;
Best Local Similarity 100.0%; Pred. No. 1.le-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 82696 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 82657

RESULT 26
AC034243/c AC034243 DNA HTG 05-APR-2000
LOCUS Homo sapiens chromosome 5 clone RPI-204011, LOW-PASS SEQUENCE
DEFINITION SAMPLING.
AC034243
ACCESSION AC034243
VERSION AC034243.1 GI:7417720

KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 166947)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
COMMENT Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

* NOTE: This record contains 191 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 572: contig of 572 bp in length
* gap of unknown length
* 573 1166: contig of 594 bp in length
* gap of unknown length
* 1167 1829: contig of 663 bp in length
* gap of unknown length
* 1830 2234: contig of 405 bp in length
* gap of unknown length
* 2235 2359: contig of 125 bp in length
* gap of unknown length
* 2360 2960: contig of 601 bp in length
* gap of unknown length
* 2961 3545: contig of 585 bp in length
* gap of unknown length
* 3546 3915: contig of 370 bp in length
* gap of unknown length
* 3916 4482: contig of 567 bp in length
* gap of unknown length
* 4483 5243: contig of 761 bp in length
* gap of unknown length
* 5244 5861: contig of 618 bp in length
* gap of unknown length
* 5862 6572: contig of 711 bp in length
* gap of unknown length
* 6573 7297: contig of 725 bp in length
* gap of unknown length
* 7298 7450: contig of 153 bp in length
* gap of unknown length
* 7451 8131: contig of 681 bp in length
* gap of unknown length
* 8132 8552: contig of 421 bp in length
* gap of unknown length
* 8553 9149: contig of 597 bp in length
* gap of unknown length
* 9150 9943: contig of 794 bp in length
* gap of unknown length
* 9944 10502: contig of 559 bp in length
* gap of unknown length
* 10503 10878: contig of 376 bp in length
* gap of unknown length
* 10879 11524: contig of 646 bp in length
* gap of unknown length
* 11525 11885: contig of 361 bp in length

```
* 11886 gap of unknown length
* 12408: contig of 523 bp in length
* 12409 gap of unknown length
* 12555: contig of 147 bp in length
* 12556 gap of unknown length
* 12893: contig of 338 bp in length
* 12894 gap of unknown length
* 13599: contig of 706 bp in length
* 13600 gap of unknown length
* 13921: contig of 322 bp in length
* 13922 gap of unknown length
* 14064: contig of 143 bp in length
* 14065 gap of unknown length
* 14143: contig of 79 bp in length
* 14144 gap of unknown length
* 14237: contig of 94 bp in length
* 14238 gap of unknown length
* 15013: contig of 776 bp in length
* 15014 gap of unknown length
* 15170: contig of 157 bp in length
* 15171 gap of unknown length
* 15509: contig of 339 bp in length
* 15510 gap of unknown length
* 15964: contig of 455 bp in length
* 15965 gap of unknown length
* 16075: contig of 111 bp in length
* 16076 gap of unknown length
* 16863: contig of 788 bp in length
* 16864 gap of unknown length
* 16997: contig of 134 bp in length
* 16998 gap of unknown length
* 18059: contig of 1062 bp in length
* 18060 gap of unknown length
* 18338: contig of 279 bp in length
* 18339 gap of unknown length
* 19199: contig of 861 bp in length
* 19200 gap of unknown length
* 19924: contig of 725 bp in length
* 19925 gap of unknown length
* 20160: contig of 236 bp in length
* 20161 gap of unknown length
* 20502: contig of 342 bp in length
* 20503 gap of unknown length
* 20571: contig of 69 bp in length
* 20572 gap of unknown length
* 20731: contig of 160 bp in length
* 20732 gap of unknown length
* 21469: contig of 738 bp in length
* 21470 gap of unknown length
* 22147: contig of 678 bp in length
* 22148 gap of unknown length
* 22872: contig of 725 bp in length
* 22873 gap of unknown length
* 23517: contig of 645 bp in length
* 23518 gap of unknown length
* 23755: contig of 238 bp in length
* 23756 gap of unknown length
* 23988: contig of 233 bp in length
* 23989 gap of unknown length
* 25107: contig of 1119 bp in length
* 25108 gap of unknown length
* 25951: contig of 844 bp in length
* 25952 gap of unknown length
* 26688: contig of 737 bp in length
* 26689 gap of unknown length
* 27404: contig of 716 bp in length
* 27405 gap of unknown length
* 28193: contig of 789 bp in length
* 28194 gap of unknown length
* 28351: contig of 158 bp in length
* 28352 gap of unknown length
* 28645: contig of 294 bp in length
* gap of unknown length
```

```
* 28646 28724: contig of 79 bp in length
* gap of unknown length
* 28725 28826: contig of 102 bp in length
* gap of unknown length
* 28827 29502: contig of 676 bp in length
* gap of unknown length
* 29503 29591: contig of 89 bp in length
* gap of unknown length
* 29592 30298: contig of 707 bp in length
* gap of unknown length
* 30299 30650: contig of 352 bp in length
* gap of unknown length
* 30651 30714: contig of 64 bp in length
* gap of unknown length
* 30715 31230: contig of 516 bp in length
* gap of unknown length
* 31231 32154: contig of 924 bp in length
* gap of unknown length
* 32155 32707: contig of 553 bp in length
* gap of unknown length
* 32708 33389: contig of 682 bp in length
* gap of unknown length
* 33390 33534: contig of 145 bp in length
* gap of unknown length
* 33535 33772: contig of 238 bp in length
* gap of unknown length
* 33773 34130: contig of 358 bp in length
* gap of unknown length
* 34131 35110: contig of 980 bp in length
* gap of unknown length
* 35111 35721: contig of 611 bp in length
* gap of unknown length
* 35722 35795: contig of 74 bp in length
* gap of unknown length
* 35796 36304: contig of 509 bp in length
* gap of unknown length
* 36305 37345: contig of 1041 bp in length
* gap of unknown length
* 37346 37732: contig of 387 bp in length
* gap of unknown length
* 37733 37989: contig of 257 bp in length
* gap of unknown length
* 37990 38420: contig of 431 bp in length
* gap of unknown length
* 38421 39039: contig of 619 bp in length
* gap of unknown length
* 39040 39185: contig of 146 bp in length
* gap of unknown length
* 39186 39755: contig of 570 bp in length
```

```
Query Match 1.3%; Score 40; DB 73; Length 166947;
Best Local Similarity 100.0%; Pred. No. 1.le-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 166035 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 165996
```

RESULT 27

```
AC015968
LOCUS AC015968 167627 bp DNA HTG 21-APR-2000
DEFINITION Homo sapiens chromosome 7 clone RP11-133L20, WORKING DRAFT
SEQUENCE, 2 unordered pieces.
ACCESSION AC015968
VERSION AC015968.3 GI:7630858
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 167627)
AUTHORS Waterston,R.H.
```

TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 167627)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT On Apr 21, 2000 this sequence version replaced gi:6524380.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H_NH0133L20
----- Summary Statistics -----
Sequencing vector: M13; 84%
Sequencing vector: plasmid; 16%
Chemistry: Dye-primer ET; 84% of reads
Chemistry: Dye-terminator Big Dye; 16% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165708 bases at least Q40
Consensus quality: 166084 bases at least Q30
Consensus quality: 166501 bases at least Q20
Insert size: 163000; agarose-fp
Insert size: 167527; sum-of-contigs
Quality coverage: 9.28 in Q20 bases; agarose-fp
Quality coverage: 9.07 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 11767: contig of 11767 bp in length
* 11768 11867: gap of unknown length
* 11868 167627: contig of 155760 bp in length.

FEATURES

Source
1. .167627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/clone="RP11-133L20"
BASE COUNT 46371 a 38471 c 38705 g 43980 t 100 others
ORIGIN

Query Match 1.3%; Score 40; DB 78; Length 167627;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 65760 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 65799

RESULT 28

AC007465 170125 bp DNA HTG 05-JUN-1999
LOCUS Homo sapiens clone NH0575121, *** SEQUENCING IN PROGRESS ***
DEFINITION unordered pieces.

AC007465
ACCESSION AC007465.2 GI:5001471
VERSION
KEYWORDS HTG: HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 170125)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170125)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT

On Jun 5, 1999 this sequence version replaced gi:4734032.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 170125: contig of 170125 bp in length.

FEATURES

Source
1. .170125
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NH0575121"
BASE COUNT 53393 a 32219 c 31961 g 52552 t
ORIGIN

Query Match 1.3%; Score 40; DB 41; Length 170125;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 113571 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 113610

RESULT 29

AL139412 170273 bp DNA HTG 19-FEB-2000
LOCUS Homo sapiens chromosome 1 clone RP11-98G7 map q21.2-22, ***
DEFINITION SEQUENCING IN PROGRESS ***
ACCESSION AL139412.2 GI:7009514
VERSION
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
Plumb.B
1 (bases 1 to 170273)
Direct Submission
Submitted (19-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

REFERENCE

AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
COMMENT On Feb 20, 2000 this sequence version replaced gi:6996233.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in
progress and the release of this data is based on the understanding
that the sequence may change as work continues. The sequence may
be contaminated with foreign sequence from E.coli, yeast, vector,
phage etc. Order of segments is not known; 800 n's separate
segments. Contig_ID: 00007 Length: 2124bp
Contig_ID: 00019 Length: 21300bp
Contig_ID: 00130 Length: 3263bp
Contig_ID: 00270 Length: 13349bp
Contig_ID: 00664 Length: 1257bp
Contig_ID: 00693 Length: 13321bp
Contig_ID: 00699 Length: 12395bp
Contig_ID: 00740 Length: 18584bp
Contig_ID: 00912 Length: 6793bp
Contig_ID: 01146 Length: 5013bp
Contig_ID: 01496 Length: 3736bp.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

*	1	2594:	gap of	800 bp	in length
*	1215	2934:	gap of	800 bp	in length
*	2925	24234:	contig of	21300 bp	in length
*	24225	25024:	gap of	800 bp	in length
*	25025	28287:	contig of	3263 bp	in length
*	28288	29087:	gap of	800 bp	in length
*	29088	42436:	contig of	13349 bp	in length
*	42437	43236:	gap of	800 bp	in length
*	43237	44493:	contig of	1257 bp	in length
*	44494	45393:	gap of	800 bp	in length
*	45294	58614:	contig of	13321 bp	in length
*	58615	59414:	gap of	800 bp	in length
*	59415	71809:	contig of	12395 bp	in length
*	71810	72609:	gap of	800 bp	in length
*	72610	91193:	contig of	18584 bp	in length
*	91194	91993:	gap of	800 bp	in length
*	91994	159924:	contig of	67931 bp	in length
*	159925	160724:	gap of	800 bp	in length
*	160725	165737:	contig of	5013 bp	in length
*	165738	166537:	gap of	800 bp	in length
*	166538	170273:	contig of	3736 bp	in length

FEATURES SOUND

```

source
1. 117023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="q21.2-22"
/clone="RP11-98G7"
/clone_lib="RPC1-11.1"
40295 a 41235 c 40421 g 40317 t 8005 others
BASE COUNT
ORIGIN

```

```

Query Match      1.3%; Score 40; DB 32; Length 170273;
Best Local Similarity 100.0%; Pred. NO. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 65755 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 65794

```

RESULT	30
AC012312	
LOCUS	AC012312 170948 bp DNA HTG 05-APR-2000
DEFINITION	Homo sapiens chromosome 5 clone CTC-558019, WORKING DRAFT SEQUENCE, 24 unordered pieces.
ACCESSION	AC012312
VERSION	AC012312.3 GI:7417591
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 170948) DOE Joint Genome Institute.
AUTHORS	Sequencing of Human Chromosome 5 Unpublished
JOURNAL	2 (bases 1 to 170948)
REFERENCE	DOE Joint Genome Institute. Direct Submission
AUTHORS	Submitted (23-OCT-1999), Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
TITLE	On Apr 5, 2000 this sequence version replaced gi:6693123. -----Genome Center
JOURNAL	
COMMENT	

BASE COUNT 54102 a 31180 c 31320 g 54266 t 80 others
ORIGIN

Query Match 1.3%; Score 40; DB 72; Length 170948;
Best Local Similarity 100.0%; Pred. NO. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 25113 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAA 25152

RESULT 31

AC009731 172749 bp DNA PRI 02-DEC-1999
LOCUS Homo sapiens 12 BAC RP11-438N16 (Roswell Park Cancer Institute
DEFINITION Human BAC Library) complete sequence.

AC009731

VERSION AC009731.6 GI:6492472

KEYWORDS HTG.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 172749)

REFERENCE
AUTHORS Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,M., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
Holloway,C., Hosak,H., Issar,A., Jackson,L.E., Jackson,L., Jia,Y.,
Jones,M., Kelly,S., Kneitz,S., Kondejewski,N., Kong,Y., Kovar,C.,
Lau,S., Leal,B., Lee,E., Li,Z., Lichtarge,O., Liu,J., Liu,W.,
Logan,O., Lozado,R.J., Lu,J., Lucier,R., Marondel,I., Martin,R.,
Martinez,C., McLeod,M.P., Mei,G., Merscher,S., Miller,A.,
Montgomery,K.T., Morgan,M., Morris,S., Nash,S., Nelson,A.,
Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B., Paxton,S.,
Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D., Rives,M.,
Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Shim,C., Simon,M.,
Sparks,A., Stamps,A., Sucgang,R., Tabor,P., Taylor,T., Vasquez,L.,
Vinson,R., Vo,Q., Wahbah,M., Watlington,S., Weinstein,G.,
Weinstock,I.R., Williamson,A., Worley,K., Wren,J., Wrensford,G.,
Xiang,A.M., Yang,R., Yu,W., Zhou,X., Kucherlapati,R., Nelson,D. and
Gibbs,R.

DIRECT SUBMISSION

UNPUBLISHED

2 (bases 1 to 172749)

Worley,K.C.

DIRECT SUBMISSION

SUBMITTED (30-AUG-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 172749).

Worley,K.C.

DIRECT SUBMISSION

SUBMITTED (01-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 172749)

Worley,K.C.

DIRECT SUBMISSION

SUBMITTED (02-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Dec 1, 1999 this sequence version replaced gi:5870175.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only

sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 172749
Phrap values in estimate: 170704
Average error rate (BCM-Phrap estimate): 0.000441707
Fraction of Phrap values less than 40 : 30
Number of consensus changing edits: 30
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
16722 tttttgaga(n)ggagcttgc tttttgaga(t)ggagcttgc
16762 tgcagtgtg(n)gatctcgct tgcagtgtg(c)gatctcgct
16787 gcaacctcac(n)ccgggttta gcaacctcac(t)ccgggttta
19892 actagagag(n)cagtagcacc actagagag(c)cagtagcacc
23304 gcaaatggag(n)ggatgatggt gcaaatggag(g)ggatgatggt
24841 gccctgggtgg(n)aggggctagt gccctgggtgg(g)aggggctagt
24856 gctagtgtt(n)aaagtctagt gctagtgtt(g)aaagtctagt
24958 aaagagtatg(n)nnaaatgcta aaagagtatg(g)aaaatgcta
24959 agagtatgn(n)naaatgctaa agagtatgg(a)aaatgctaa
24960 agagtatgn(n)aaatgctaaa agagtatga(a)aaatgctaaa
31414 ggcgggggggg(n)tgacccccc ggcgggggggg(c)tgacccccc
31583 ggcgggggggg(n)cggggggggt ggcgggggggg(g)cggggggggt
31588 gccgncggg(n)ggngtgaccc gccgncggg(g)ggngtgaccc
31592 gncgggggg(n)tgacccccc gncgggggg(g)tgacccccc
65790 attgataaga(n)aaataagct attgataaga(a)aaataagct
67402 ataagcattg(n)ttgggacagc ataagcattg(c)ttgggacagc
73762 ggcatagccc(n)cccccccc ttccataccc(t)ttcttttt
86111 ttccataccc(n)ttcttttt taacctcttc(t)ttcttttt
86115 taacctcttc(n)ttcttttt taacctcttc(c)ctcgccctcc
90092 aatctctctg(n)ctcgnctcc tctgctctg(t)ctcggcagaag
90097 tctgctctg(n)ctcggcagaag cccaccacga(c)ctctctccc
90285 cccaccacga(n)ctctctccc aatgtggca(n)caaaatttgc
111002 aatgtggca(n)caaaatttgc aaaaacaaa(a)caaaacaga
112201 aaaaacaaa(c)caaaacaga aaaaacaaa(a)caaaacaga
112207 aaaaacaaa(n)caaaacaga acaagacaa(n)caagctattt
112710 acaagacaa(n)caagctattt aagattcaga(c)gtgccagata
160933 aagattcaga(n)gtgccagata

* 6004 7129: contig of 1126 bp in length
* gap of unknown length
* 7130 8428: contig of 1299 bp in length
* gap of unknown length
* 8429 10151: contig of 1723 bp in length
* gap of unknown length
* 10152 10364: contig of 213 bp in length
* gap of unknown length
* 10365 11828: contig of 1464 bp in length
* gap of unknown length
* 11829 13207: contig of 1379 bp in length
* gap of unknown length
* 13208 14794: contig of 1587 bp in length
* gap of unknown length
* 14795 16552: contig of 1758 bp in length
* gap of unknown length
* 16553 18749: contig of 2197 bp in length
* gap of unknown length
* 18750 21181: contig of 2432 bp in length
* gap of unknown length
* 21182 23285: contig of 2104 bp in length
* gap of unknown length
* 23286 25187: contig of 1902 bp in length
* gap of unknown length
* 25188 27119: contig of 1932 bp in length
* gap of unknown length
* 27120 29572: contig of 2453 bp in length
* gap of unknown length
* 29573 31887: contig of 2315 bp in length
* gap of unknown length
* 31888 34036: contig of 2149 bp in length
* gap of unknown length
* 34037 35754: contig of 1718 bp in length
* gap of unknown length
* 35755 38353: contig of 2599 bp in length
* gap of unknown length
* 38354 41887: contig of 3534 bp in length
* gap of unknown length
* 41888 44323: contig of 2436 bp in length
* gap of unknown length
* 44324 46990: contig of 2667 bp in length
* gap of unknown length
* 46991 50211: contig of 3221 bp in length
* gap of unknown length
* 50212 53782: contig of 3571 bp in length
* gap of unknown length
* 53783 57135: contig of 3353 bp in length
* gap of unknown length
* 57136 60737: contig of 3602 bp in length
* gap of unknown length
* 60738 64425: contig of 3688 bp in length
* gap of unknown length
* 64426 69302: contig of 4877 bp in length
* gap of unknown length
* 69303 74567: contig of 5265 bp in length
* gap of unknown length
* 74568 79311: contig of 4744 bp in length
* gap of unknown length
* 79312 82695: contig of 3384 bp in length
* gap of unknown length
* 82696 88579: contig of 5884 bp in length
* gap of unknown length
* 88580 93341: contig of 4762 bp in length
* gap of unknown length
* 93342 97484: contig of 4143 bp in length
* gap of unknown length
* 97485 101749: contig of 4265 bp in length
* gap of unknown length
* 101750 108833: contig of 7084 bp in length
* gap of unknown length
* 108834 114784: contig of 5951 bp in length
* gap of unknown length
* 114785 123180: contig of 8396 bp in length

* 123181 129628: contig of 6448 bp in length
* gap of unknown length
* 129629 137374: contig of 7746 bp in length
* gap of unknown length
* 137375 145275: contig of 7901 bp in length
* gap of unknown length
* 145276 153859: contig of 8584 bp in length
* gap of unknown length
* 153860 165597: contig of 11738 bp in length
* gap of unknown length
* 165598 176771: contig of 11174 bp in length
* gap of unknown length
* 176772 193632: contig of 16861 bp in length.

FEATURES

source

1. .193632
/organism="Homo sapiens"
/db_xref="taxon:9606"
/Chromosome="15"
/map="15"
/clone="RP11-344C1"
/clone_lib="RPC1-11 Human Male BAC"
BASE COUNT 57762 a 38404 c 38424 g 59032 t 10 others
ORIGIN

Query Match 1.3%; Score 40; DB 43; Length 193632;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 55291 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 55252

RESULT 33

HS80N2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

HS80N2 193731 bp DNA PRI 02-MAR-2000
Human DNA sequence from clone RP1-80N2 on chromosome 6p24.1-25.3.
Contains the gene for MD-1, part of a putative novel gene, ESTs,
STSs, GSSs and three putative CpG islands, complete sequence.
AL031123
AL031123.14 GI:6911574
HTG: Cpg island; MD-1.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Williams, S.
Direct Submission
Submitted (18-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Feb 7, 2000 this sequence version replaced gi:6691904.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human

chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>
RPI-80N2 is from the library RPCI-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:PCYPAC2>
This sequence is the entire insert of clone RPI-80N2.

FEATURES

source	Location/Qualifiers	repeat_region	repeat_region
	1..193731		16317..16626
	/organism="Homo sapiens"		/note="AluX repeat: matches 1. .311 of consensus"
	/db_xref="taxon:9606"	repeat_region	16629..16893
	/chromosome="6"		/note="AluX repeat: matches 2. .302 of consensus"
	/map="p24.1-25.3"	misc_feature	16713..17267
	/clone="RPI-80N2"	repeat_region	/note="match: GSS: Em:AQ791339"
	/clone_lib="RPC1-1"	repeat_region	20655..20731
	202..490	repeat_region	/note="L1MD1 repeat: matches 6077. .6167 of consensus"
	/note="AluJb repeat: matches 3. .277 of consensus"	repeat_region	20667..21229
	494..609	repeat_region	/note="L1MD2 repeat: matches 5679. .6269 of consensus"
	/note="29 copies 4 mer agag 65% conserved"	repeat_region	21245..22937
	1628..1694	repeat_region	/note="L1MC1 repeat: matches 3566. .5403 of consensus"
	/note="MIR repeat: matches 64. .131 of consensus"	repeat_region	22942..23535
	1958..2553	repeat_region	/note="MER68A repeat: matches 3. .577 of consensus"
	/note="L2 repeat: matches 1630. .2278 of consensus"	repeat_region	23539..23639
	2655..3074	repeat_region	/note="L1MC1 repeat: matches 3446. .3542 of consensus"
	3081..3458	repeat_region	23640..23724
	/note="L1M4 repeat: matches 3519. .3950 of consensus"	repeat_region	/note="MER44C repeat: matches 1. .85 of consensus"
	3459..3667	repeat_region	23728..24229
	/note="AluJo repeat: matches 88. .295 of consensus"	repeat_region	/note="MER44C repeat: matches 207. .728 of consensus"
	3673..4228	repeat_region	24230..24442
	/note="L1M4 repeat: matches 2937. .3516 of consensus"	repeat_region	/note="L1MC1 repeat: matches 3231. .3447 of consensus"
	4933..5727	repeat_region	24443..24679
	/note="L1MC1 repeat: matches 5560. .6332 of consensus"	repeat_region	/note="AluY repeat: matches 75. .311 of consensus"
	6975..7261	repeat_region	24680..25198
	/note="AluSg repeat: matches 1. .293 of consensus"	repeat_region	/note="L1MC1 repeat: matches 2704. .3231 of consensus"
	7672..7936	repeat_region	25199..25572
	/note="AluSg repeat: matches 42. .300 of consensus"	repeat_region	/note="MSTA repeat: matches 1. .426 of consensus"
	8517..8663	repeat_region	25573..27155
	/note="AluSq/x repeat: matches 2. .148 of consensus"	repeat_region	/note="MSTA-internal repeat: matches 3. .1580 of consensus"
	9034..9302	misc_feature	27158..27538
	/note="match: GSS: Em:AQ277477"	misc_feature	27539..27739
	9082..9941		/note="L1MC1 repeat: matches 2527. .2704 of consensus"
	/note="CpG island"		29283..29421
	/evidence=not_experimental		/note="L2 repeat: matches 2544. .2697 of consensus"
	11272..11501	repeat_region	29539..29640
	/note="HALI repeat: matches 150. .394 of consensus"	repeat_region	/note="MER3 repeat: matches 1. .103 of consensus"
	11504..11551	repeat_region	29713..29863
	/note="MER45 repeat: matches 7. .54 of consensus"	repeat_region	/note="MER5A repeat: matches 3. .188 of consensus"
	11564..11849	repeat_region	30005..30463
	/note="AluSc repeat: matches 1. .285 of consensus"	repeat_region	/note="MER69 repeat: matches 214. .717 of consensus"
	11868..12079	repeat_region	30486..30730
	/note="AluJo repeat: matches 20. .230 of consensus"	repeat_region	/note="L1PA16 repeat: matches 5889. .6157 of consensus"
	12100..12350	repeat_region	30747..31274
	/note="AluJo repeat: matches 1. .291 of consensus"	repeat_region	/note="MER69 repeat: matches 1954. .2510 of consensus"
	12399..12461	repeat_region	31294..31420
	/note="AluJo repeat: matches 226. .287 of consensus"	repeat_region	/note="L2 repeat: matches 2629. .2749 of consensus"
	12459..12513	repeat_region	33424..33695
	/note="MER45B repeat: matches 924. .978 of consensus"	repeat_region	/note="AluSc repeat: matches 37. .309 of consensus"
	12704..12974	repeat_region	33696..33743
	/note="MER33 repeat: matches 14. .307 of consensus"	repeat_region	/note="L2 copies 4 mer gga 95% conserved"
	13684..13952	repeat_region	34330..36441
	/note="HALI repeat: matches 1128. .1412 of consensus"	repeat_region	/note="L1MB5 repeat: matches 4015. .6138 of consensus"
	complement(13927..14420)	misc_feature	complement(36498..37004)
	/note="match: GSS: Em:AQ110046"		/note="match: GSS: Em:AQ833420"
	complement(14008..14404)		36561..36828
	/note="match: GSS: Em:AQ040995"		/note="L1MEC repeat: matches 271. .558 of consensus"
	14695..15038	repeat_region	37241..38195
	/note="MLR1I repeat: matches 79. .403 of consensus"	repeat_region	/note="L1MEC repeat: matches 1025. .2103 of consensus"
	15245..15596	repeat_region	38402..38696
	/note="MLR1A1 repeat: matches 1. .365 of consensus"	repeat_region	/note="L1R40a repeat: matches 217. .519 of consensus"
	15934..16038	repeat_region	38701..38827
	/note="AluY repeat: matches 196. .300 of consensus"	repeat_region	/note="L1MEC repeat: matches 2102. .2230 of consensus"
		repeat_region	38999..39314
		repeat_region	/note="MER33 repeat: matches 2. .324 of consensus"
		repeat_region	39315..39450
		repeat_region	/note="L1MEC repeat: matches 2296. .2093 of consensus"
		repeat_region	39669..39882
		repeat_region	/note="MIR repeat: matches 2. .262 of consensus"
		repeat_region	40576..40687
		repeat_region	/note="28 copies 4 mer tata 77% conserved"
		repeat_region	40938..41122

```
/note="AluJo repeat: matches 122. .299 of consensus"
42029. .42097
/note="L1MC/D repeat: matches 5752. .5819 of consensus"
42167. .42204
/note="19 copies 2 mer gt 89% conserved"
42170. .42205
/note="9 copies 4 mer tgtg 88% conserved"
43073. .43153
/note="L2 repeat: matches 2067. .2146 of consensus"
43193. .43528
/note="THE1B repeat: matches 1. .364 of consensus"
43600. .43907
/note="AluJo repeat: matches 5. .310 of consensus"
44113. .44399
/note="STS: Em:Z79002"
44859. .44900
/note="L1M3c repeat: matches 273. .314 of consensus"
44915. .45163

Query Match 1.3%; Score 40; DB 11; Length 193731;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tcatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 63526 TCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 63565

RESULT 34
BTClB8 583 bp mRNA MAM 27-AUG-1992
LOCUS B.taurus CI-B8 mRNA for ubiquinone oxidoreductase complex.
DEFINITION X63219
ACCESSION X63219
VERSION X63219.1 GI:245
KEYWORDS NADH dehydrogenase; NADH-ubiquinone oxidoreductase complex.
SOURCE Bos taurus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 583)
Walker,J.
Direct Submission
Submitted (12-NOV-1991) Walker J., MRC Lab. of Molecular Biology,
Hills Road, Cambridge CB2 2QH, England
2 (bases 1 to 583)
Walker,J.E., Arizmendi,J.M., Dupuis,A., Fearnley,I.M., Finel,M.,
Medd,S.M., Pilkington,S.J., Runswick,M.J. and Skehel,J.M.
Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from
bovine heart mitochondria. Application of a novel strategy for
sequencing proteins using the polymerase chain reaction
J. Mol. Biol. 226 (4), 1051-1072 (1992)
92389317 Location/Qualifiers
FEATURES
1. .583
/organism="Bos taurus"
/db_xref="taxon:9913"
/germline
/tissue_type="heart"
34. .333
/gene="CI-B8"
34. .333
/gene="CI-B8"
/EC_number="1.6.99.3"
/codon_start=1
evidence=experimental
/product="NADH dehydrogenase"
/protein_id="CAA44904.1"
/db_xref="GI:246"
/db_xref="SWISS-PROT:Q02370"
/translation="MAAAATRGVKGKGLRIRIHLQSRPSGSGVDFIKRVEL
KKANPLFILRECSVQPKLWARYAFQGNVSLNFSADQVTRALENVLSKA"
523. .528
polyA_signal
polyA_site
```

```
/evidence=experimental
183 a 126 c 143 g 131 t
BASE COUNT
ORIGIN

Query Match 1.3%; Score 39; DB 3; Length 583;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 545 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 583

RESULT 35
AF206698 606 bp mRNA INV 09-JAN-2000
LOCUS Epiblena scudderiana muscle LIM protein (Mlp) mRNA, complete cds.
DEFINITION AF206698
ACCESSION AF206698
VERSION AF206698.1 GI:6683944
KEYWORDS Epiblena scudderiana.
SOURCE Epiblena scudderiana.
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Tortricodea; Tortricidae; Olethreutinae; Epiblema.
1 (bases 1 to 606)
Bilgen,T., English,T.E. and Storey,K.B.
EsMlp, a muscle-LIM protein gene, is up-regulated during cold
exposure in the freeze-avoiding larvae of Epiblema scudderiana
Unpublished
2 (bases 1 to 606)
Bilgen,T. and Storey,K.B.
Direct Submission
Submitted (18-NOV-1999) Biology, Carleton University, 1125 Colonel
By Drive, Ottawa, ON K1S 5B6, Canada
Location/Qualifiers
1. .606
/organism="Epiblema scudderiana"
/db_xref="taxon:111441"
/note="induced by subzero exposure; highest expression
apparently localized in pharyngeal muscles"
1. .606
/gene="Mlp"
93. .377
/gene="Mlp"
/codon_start=1
/product="muscle LIM protein"
/protein_id="AAF23406.1"
/db_xref="GI:6683945"
/translation="MPFKPADNPKPCGKSVAAERVAGGLKWHKTCFKGCMCNKS
LDSTNCTEHDGEIYCKNCHGRFKGPKGYGGGACLSMDAQAQENAN"
192 a 153 c 141 g 120 t
BASE COUNT
ORIGIN

Query Match 1.3%; Score 39; DB 34; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 565 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 603

RESULT 36
AF123275 1845 bp DNA INV 29-MAR-1999
LOCUS Dictyostellium discoideum superoxide-generating NADPH oxidase
DEFINITION AF123275
ACCESSION AF123275
VERSION AF123275.1 GI:4530485
KEYWORDS
```

```

SOURCE Dictyostellium discoideum.
ORGANISM Dictyostellium discoideum
REFERENCE Dictyostellida; Dictyostellum.
AUTHORS 1 (bases 1 to 1845)
TITLE Bof. M. and Satre.M.
Nucleotide sequence of Dictyostellium discoideum
superoxide-generating NADPH oxidase heavy chain
Unpublished
JOURNAL 2 (bases 1 to 1845)
AUTHORS Bof. M. and Satre.M.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-1999) DBMS/BBSI (UMR 314 CNRS), CEA-Grenoble, 17
rue des Martyrs, 38054 Grenoble Cedex 9, France
FEATURES
    source
        1..1845
            Location/Qualifiers
                /organism="Dictyostellium discoideum"
                /strain="Ax-2"
                /db_xref="taxon:44689"
                join(<1..118,229..457,639..>.1845)
                /product="superoxide-generating NADPH oxidase
                flavocytochrome"
                join(1..118,229..457,639..1845)
                /note="gp91phox"
                /codon_start=1
                /product="superoxide-generating NADPH oxidase
                flavocytochrome"
                /protein_id="AAD22057.1"
                /db_xref="GI:4530486"
                /translation="MRLPKEIQRYWYNEGKLLILVILYLTGNTAAFYVTVHVVYNS
                PAFEVGVGFARGCAQLKLNLCALILVPLRNLISFLRGFTLNNYVFPDKNIYFHK
                LIAWICFATGTHVNAHFNFRFYQDITPOEYKRILGIDYPLNLTPIKYAFATLAGWTG
                HVVICVLMVLTSAVESIRSRPMFEGFWTHLFVVFGLVVLVHGLHSILEPTSFWKWV
                IGPCALYVERILIRLLRSKKTMLIQARHPISRVTEVRMKTFRFKYKPCOYLFLNCP
                IAQNEHPHTIRISAPEDFVSCHINVGNWGKLTLLNPKMKMGIOENVLKSDDGK
                FILRIDFGAASEVFYKQVILVAGIGVTFPASILKHLNKIQWARTINTPLLDIKV
                HFYICWRNRSFENFWSGLIELEMNHNHNFLEIHPYLTGALSAQEIIRDVDMYGDGEKOL
                ITGTFPTQFGRPKDEIFADHALRYAEKDVGFVFCGPKLLSKSLYKASTHYTKTTTC
                RFHYKNENF"
BASE COUNT 665 a 226 c 275 g 679 t
ORIGIN

Query Match 1.3%; Score 39; DB 34; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 149 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 187

RESULT 37
HUMSODIUM 2334 bp DNA PRI 29-SEP-1994
LOCUS Human kidney amiloride-sensitive sodium channel, complete cds.
DEFINITION
ACCESSION L29007
VERSION L29007 1 GI:493125
KEYWORDS sodium channel.
SOURCE Homo sapiens Kidney DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2334)
TITLE McDonald, F.J., Snyder, P.M., McCray, P.B. Jr. and Welsh, M.J.
Cloning, expression, and tissue distribution of a human
amiloride-sensitive Na+ channel
JOURNAL Am. J. Physiol. 266 (Pt 1), L728-L734 (1994)
MEDLINE 94295729
FEATURES
    source
        1..2334
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /tissue_type="Kidney"

SOURCE Dictyostellium discoideum.
ORGANISM Dictyostellium discoideum
REFERENCE Dictyostellida; Dictyostellum.
AUTHORS 1 (bases 1 to 1845)
TITLE Bof. M. and Satre.M.
Nucleotide sequence of Dictyostellium discoideum
superoxide-generating NADPH oxidase heavy chain
Unpublished
JOURNAL 2 (bases 1 to 1845)
AUTHORS Bof. M. and Satre.M.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-1999) DBMS/BBSI (UMR 314 CNRS), CEA-Grenoble, 17
rue des Martyrs, 38054 Grenoble Cedex 9, France
FEATURES
    source
        1..1845
            Location/Qualifiers
                /organism="Dictyostellium discoideum"
                /strain="Ax-2"
                /db_xref="taxon:44689"
                join(<1..118,229..457,639..>.1845)
                /product="superoxide-generating NADPH oxidase
                flavocytochrome"
                join(1..118,229..457,639..1845)
                /note="gp91phox"
                /codon_start=1
                /product="superoxide-generating NADPH oxidase
                flavocytochrome"
                /protein_id="AAD22057.1"
                /db_xref="GI:4530486"
                /translation="MRLPKEIQRYWYNEGKLLILVILYLTGNTAAFYVTVHVVYNS
                PAFEVGVGFARGCAQLKLNLCALILVPLRNLISFLRGFTLNNYVFPDKNIYFHK
                LIAWICFATGTHVNAHFNFRFYQDITPOEYKRILGIDYPLNLTPIKYAFATLAGWTG
                HVVICVLMVLTSAVESIRSRPMFEGFWTHLFVVFGLVVLVHGLHSILEPTSFWKWV
                IGPCALYVERILIRLLRSKKTMLIQARHPISRVTEVRMKTFRFKYKPCOYLFLNCP
                IAQNEHPHTIRISAPEDFVSCHINVGNWGKLTLLNPKMKMGIOENVLKSDDGK
                FILRIDFGAASEVFYKQVILVAGIGVTFPASILKHLNKIQWARTINTPLLDIKV
                HFYICWRNRSFENFWSGLIELEMNHNHNFLEIHPYLTGALSAQEIIRDVDMYGDGEKOL
                ITGTFPTQFGRPKDEIFADHALRYAEKDVGFVFCGPKLLSKSLYKASTHYTKTTTC
                RFHYKNENF"
BASE COUNT 665 a 226 c 275 g 679 t
ORIGIN

Query Match 1.3%; Score 39; DB 9; Length 2334;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcctgtccaaacctggagccagtcgagccctgtggtt 1838
|||||
DB 1720 CCTCCTGTCCAACTGGGCGCAGTCGAGCCGTGTGGTT 1758

RESULT 38
HSAENAC4 2562 bp DNA PRI 02-MAY-1999
LOCUS Homo sapiens epithelial sodium channel alpha-subunit gene, exons 9
DEFINITION
ACCESSION AF060913
VERSION AF060913.1 GI:4731107
KEYWORDS
SEGMENT 4 of 4
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2562)
TITLE Chow, Y.H., Wang, Y., Plumb, J., O'Brodovich, H. and Hu, J.
Hormonal regulation and genomic organization of the human
amiloride-sensitive epithelial sodium channel alpha-subunit gene
JOURNAL Pediatr. Res. (1999) In press
REFERENCE 2 (bases 1 to 2562)
AUTHORS Chow, Y.H. and Hu, J.
TITLE Direct Submission
JOURNAL Submitted (22-APR-1998) Lung Biology, Hospital for Sick Children,
555 University Ave, Toronto, ON M5G 1X8, Canada
FEATURES
    source
        1..2562
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="12"
                /map="12p13"
                join(AF060910.1:3085..3129,AF060910.1:3797..4266,
                AF060911.1:1675..1942,AF060911.1:3781..3971,
                AF060912.1:1728..831,AF060912.1:1173..1336,
                AF060912.1:1760..1858,AF060912.1:2053..2170,20..98,
                204..261,420..475,623..698,1172..2562)
                /product="epithelial sodium channel alpha-subunit"
                join(AF060910.1:3851..4266,AF060911.1:1675..1942,
                AF060911.1:3781..3971,AF060912.1:1728..831,
                AF060912.1:1173..1336,AF060912.1:1760..1858,
                AF060912.1:2053..2170,20..98,204..261,420..475,623..698,
                1172..1552)
                /note="alpha-hENaC"
                /codon_start=1

SOURCE Dictyostellium discoideum.
ORGANISM Dictyostellium discoideum
REFERENCE Dictyostellida; Dictyostellum.
AUTHORS 1 (bases 1 to 1845)
TITLE Bof. M. and Satre.M.
Nucleotide sequence of Dictyostellium discoideum
superoxide-generating NADPH oxidase heavy chain
Unpublished
JOURNAL 2 (bases 1 to 1845)
AUTHORS Bof. M. and Satre.M.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-1999) DBMS/BBSI (UMR 314 CNRS), CEA-Grenoble, 17
rue des Martyrs, 38054 Grenoble Cedex 9, France
FEATURES
    source
        1..1845
            Location/Qualifiers
                /organism="Dictyostellium discoideum"
                /strain="Ax-2"
                /db_xref="taxon:44689"
                join(<1..118,229..457,639..>.1845)
                /product="superoxide-generating NADPH oxidase
                flavocytochrome"
                join(1..118,229..457,639..1845)
                /note="gp91phox"
                /codon_start=1
                /product="superoxide-generating NADPH oxidase
                flavocytochrome"
                /protein_id="AAD22057.1"
                /db_xref="GI:4530486"
                /translation="MRLPKEIQRYWYNEGKLLILVILYLTGNTAAFYVTVHVVYNS
                PAFEVGVGFARGCAQLKLNLCALILVPLRNLISFLRGFTLNNYVFPDKNIYFHK
                LIAWICFATGTHVNAHFNFRFYQDITPOEYKRILGIDYPLNLTPIKYAFATLAGWTG
                HVVICVLMVLTSAVESIRSRPMFEGFWTHLFVVFGLVVLVHGLHSILEPTSFWKWV
                IGPCALYVERILIRLLRSKKTMLIQARHPISRVTEVRMKTFRFKYKPCOYLFLNCP
                IAQNEHPHTIRISAPEDFVSCHINVGNWGKLTLLNPKMKMGIOENVLKSDDGK
                FILRIDFGAASEVFYKQVILVAGIGVTFPASILKHLNKIQWARTINTPLLDIKV
                HFYICWRNRSFENFWSGLIELEMNHNHNFLEIHPYLTGALSAQEIIRDVDMYGDGEKOL
                ITGTFPTQFGRPKDEIFADHALRYAEKDVGFVFCGPKLLSKSLYKASTHYTKTTTC
                RFHYKNENF"
BASE COUNT 665 a 226 c 275 g 679 t
ORIGIN

Query Match 1.3%; Score 39; DB 9; Length 2334;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcctgtccaaacctggagccagtcgagccctgtggtt 1838
|||||
DB 1720 CCTCCTGTCCAACTGGGCGCAGTCGAGCCGTGTGGTT 1758

RESULT 38
HSAENAC4 2562 bp DNA PRI 02-MAY-1999
LOCUS Homo sapiens epithelial sodium channel alpha-subunit gene, exons 9
DEFINITION
ACCESSION AF060913
VERSION AF060913.1 GI:4731107
KEYWORDS
SEGMENT 4 of 4
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2562)
TITLE Chow, Y.H., Wang, Y., Plumb, J., O'Brodovich, H. and Hu, J.
Hormonal regulation and genomic organization of the human
amiloride-sensitive epithelial sodium channel alpha-subunit gene
JOURNAL Pediatr. Res. (1999) In press
REFERENCE 2 (bases 1 to 2562)
AUTHORS Chow, Y.H. and Hu, J.
TITLE Direct Submission
JOURNAL Submitted (22-APR-1998) Lung Biology, Hospital for Sick Children,
555 University Ave, Toronto, ON M5G 1X8, Canada
FEATURES
    source
        1..2562
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="12"
                /map="12p13"
                join(AF060910.1:3085..3129,AF060910.1:3797..4266,
                AF060911.1:1675..1942,AF060911.1:3781..3971,
                AF060912.1:1728..831,AF060912.1:1173..1336,
                AF060912.1:1760..1858,AF060912.1:2053..2170,20..98,
                204..261,420..475,623..698,1172..2562)
                /product="epithelial sodium channel alpha-subunit"
                join(AF060910.1:3851..4266,AF060911.1:1675..1942,
                AF060911.1:3781..3971,AF060912.1:1728..831,
                AF060912.1:1173..1336,AF060912.1:1760..1858,
                AF060912.1:2053..2170,20..98,204..261,420..475,623..698,
                1172..1552)
                /note="alpha-hENaC"
                /codon_start=1
```

```
/product="epithelial sodium channel alpha-subunit"
/db_xref="AAD28355.1"
/translation="MEGNKLEQDSSPQSTPGLMKGNKREOGLGPPEAPAOQPTAE
GLLFEHRSYRELFEFFCNNTIIGAIRLVCQSHNRMTAFWVWLCTFGMMYQWF
GLYGEISYPSVLNLSNLSKGLTLPVAVTICTLNPYRPEIKEELEELDRITEQTUFD
LYKYSFTTLVAGSRSRDLRGLTLPQLRVRPPPHGARRARSVASSLRDNNQVOD
WKDMKIGFOLCQNKSDCFYQYSSGVDAVREWYRPHYINILSRLEPTLPSLEEDTLG
NFIFACRNOVSCNANYSHFHPMGVNCYTFENDKNSNLMWSSMPGINNGLSLMLRA
EONDRIPLLSTVIGARVMVHGOPEAFMDGDFNLPGVETISMRKETLDRLGDDYG
DCTKNGSDVPVENLPSKYTQQVCIHSCFQESMIKECCGAYIFYPRPNVEYCDYRKH
SWGKYCYLQVDFSDHLGCTCKRKCSTVTSYQLSAGISRWPSVTSQEWYFQMLSR
QNNYVNNKRGVAKVNIFFKELNKTNSPSTVMTLLSLNGSQWLSWFGSSVLSV
VMAELVFDLLVIMFLMLLRFRSRYWSPGRGGAQEVASTLASSPPSHFCHPMSL
SLSQGPAPSPALTAPPAYATLGPSPGSGAGASSSTCPLGGP"
20..98
/number=9
exon
204..261
/number=10
exon
420..475
/number=11
exon
623..698
/number=12
exon
1172..2562
/number=13
exon
BASE COUNT 533 a 747 c 689 g 593 t
ORIGIN

Query Match 1.3%; Score 39; DB 39; Length 2562;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcctgtccaaacctggcagcagtgagcctgtggtt 1838
|||||
Db 1180 CCTCCTGTCCAACTGGCAGTGAGCGAGTGGAGCTGTGTT 1218

RESULT 39
HSLASNA HSLASNA 3151 bp mRNA PRI 01-FEB-1994
LOCUS H.sapiens mRNA for lung amiloride sensitive Na+ channel protein.
DEFINITION X76180
ACCESSION X76180
VERSION 1
KEYWORDS Na+ channel; Na+ channel protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3151)
Barbry, P.
Direct Submission
Submitted (19-NOV-1993) P. Barbry, CNRS, IPCM 660 Route des
Lucioles, 06560 Sophia Antipolis, FRANCE
2 (bases 1 to 3151)
Voilley, N., Lingueglia, E., Champigny, G., Mattei, M.G., Waldmann, R.,
Lazdunski, M. and Barbry, P.
The lung amiloride-sensitive Na+ channel: biophysical properties,
pharmacology, ontogenesis, and molecular cloning
Proc. Natl. Acad. Sci. U.S.A. 91 (1), 247-251 (1994)
94105144
FEATURES
Location/Qualifiers
1..3151
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/tissue_type="lung"
/cell_type="epithelial"
/dev_stage="adult"
100..2109
/codon_start=1
/product="Na+ channel protein"
/protein_id="CAA53773.1"

CDS
```

```
/db_xref="GI:452650"
/translation="MEGNKLEQDSSPQSTPGLMKGNKREOGLGPPEAPAOQPTAE
EALIEFHRSYRELFEFFCNNTIIGAIRLVCQSHNRMTAFWVWLCTFGMMYQWF
GLYGEISYPSVLNLSNLSKGLTLPVAVTICTLNPYRPEIKEELEELDRITEQTUFD
LYKYSFTTLVAGSRSRDLRGLTLPQLRVRPPPHGARRARSVASSLRDNNQVOD
WKDMKIGFOLCQNKSDCFYQYSSGVDAVREWYRPHYINILSRLEPTLPSLEEDTLG
NFIFACRNOVSCNANYSHFHPMGVNCYTFENDKNSNLMWSSMPGINNGLSLMLRA
EONDRIPLLSTVIGARVMVHGOPEAFMDGDFNLPGVETISMRKETLDRLGDDYG
DCTKNGSDVPVENLPSKYTQQVCIHSCFQESMIKECCGAYIFYPRPNVEYCDYRKH
SWGKYCYLQVDFSDHLGCTCKRKCSTVTSYQLSAGISRWPSVTSQEWYFQMLSR
QNNYVNNKRGVAKVNIFFKELNKTNSPSTVMTLLSLNGSQWLSWFGSSVLSV
VMAELVFDLLVIMFLMLLRFRSRYWSPGRGGAQEVASTLASSPPSHFCHPMSL
SLSQGPAPSPALTAPPAYATLGPSPGSGAGASSSTCPLGGP"
677 a 995 c 803 g 676 t
ORIGIN

Query Match 1.3%; Score 39; DB 9; Length 3151;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcctgtccaaacctggcagcagtgagcctgtggtt 1838
|||||
Db 1737 CCTCCTGTCCAACTGGCAGTGAGCGAGTGGAGCTGTGTT 1775

RESULT 40
A39975 A39975 4028 bp DNA PAT 05-MAR-1997
LOCUS Sequence 8 from Patent WO9421790.
DEFINITION A39975
ACCESSION A39975
VERSION A39975.1 GI:2296171
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4028)
AUTHORS Hayden, M., Goldberg, P., Andrew, S., Rommens, J.M., Lin and Blaooyang,
TITLE PROCESS FOR ISOLATING GENES AND THE GENE CAUSATIVE OF HUNTINGTON'S
JOURNAL DISEASE AND DIFFERENTIAL 3, POLYADENYLATION IN THE GENE
Patent: WO 9421790-A 8 29-SEP-1994;
UNIV BRITISH COLUMBIA (CA)
COMMENT Other publication CA 2092455 940926
Other publication AU 6281094 941011
Other publication US 5534438 960709
Other publication CA 2158918 940929.
FEATURES
Location/Qualifiers
1..4028
/organism="unidentified"
/db_xref="taxon:32644"
/chromosome="4"
BASE COUNT 838 a 1100 c 1140 g 950 t
ORIGIN

Query Match 1.3%; Score 39; DB 5; Length 4028;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 3984 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4022

RESULT 41
I23502 I23502 4032 bp DNA PAT 07-OCT-1996
LOCUS Sequence 8 from patent US 5534438.
DEFINITION I23502
ACCESSION I23502
VERSION I23502.1 GI:1603372
KEYWORDS Unknown.
SOURCE
```



```
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4032)
AUTHORS Hayden,M., Goldberg,P., Andrew,S. and Rommens,J.M.
TITLE Process for isolating genes and the gene causative of Huntington's
JOURNAL disease and differential 3' polyadenylation in the gene
FEATURES Patent: US 5534438-A 8 09-JUL-1996;
source Location/Qualifiers
1..4032
/organism="unknown"
BASE COUNT 840 a 1100 c 1142 g 950 t
ORIGIN

Query Match 1.3%; Score 39; DB 5; Length 4032;
Best Local Similarity 100.0%; Pred.No.3e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 3988 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4026

RESULT 42
HSZ92981
LOCUS HSZ92981 6313 bp DNA PRI 13-JUL-1998
DEFINITION H sapiens SCNN1A gene, exons 5-13.
ACCESSION Z92981
VERSION Z92981.1 GI:2765705
KEYWORDS epithelial amiloride-sensitive sodium channel; SCNN1A gene; sodium
channel alpha subunit.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6313)
AUTHORS Ludwig,M.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1997) Ludwig M., University of Bonn, Dep. of
Clinical Biochemistry, Sigmund-Freud-Strasse 25, Bonn, NRW,
Germany, 53105
REFERENCE 2 (bases 1 to 6313)
AUTHORS Ludwig,M., Bolkenius,U., Wickert,L., Marynen,P. and Bidlingmaier,F.
TITLE Structural organisation of the gene encoding the alpha-subunit of
the human amiloride-sensitive epithelial sodium channel
JOURNAL Hum. Genet. 102 (5), 576-581 (1998)
MEDLINE 98316780
FEATURES Location/Qualifiers
source 1..6313
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosomes="12"
<1..1668
/citation=[2]
/number=4
repeat_unit <1..261
/rpt_family="Alu"
gene 1669..6313
/organism="SCNN1A"
exon 1669..1772
/organism="SCNN1A"
/standard_name="ENaC"
/number=5
/usedin=Z92978:SCNN1A_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
intron 1773..2114
/organism="SCNN1A"
/number=5
exon 2115..2278
/organism="SCNN1A"
/standard_name="ENaC"
/number=6

/usedin=Z92978:SCNN1A_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
intron 2279..2702
/organism="SCNN1A"
/number=6
exon 2703..2801
/organism="SCNN1A"
/standard_name="ENaC"
/number=7
/usedin=Z92978:SCNN1A_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
intron 2802..2995
/organism="SCNN1A"
/number=7
exon 2996..3113
/organism="SCNN1A"
/standard_name="ENaC"
/number=8
/usedin=Z92978:SCNN1A_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
intron 3114..3755
/organism="SCNN1A"
/number=8
repeat_unit 3316..3571
/organism="SCNN1A"
/rpt_family="Alu"
exon 3756..3834
/organism="SCNN1A"
/standard_name="ENaC"
/number=9
/usedin=Z92978:SCNN1A_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
intron 3835..3939
/organism="SCNN1A"
/number=9
exon 3940..3997
/organism="SCNN1A"
/standard_name="ENaC"
/number=10
/usedin=Z92978:SCNN1A_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
intron 3998..4155
/organism="SCNN1A"
/number=10
intron 4112..4358
/organism="SCNN1A"
/number=11
exon 4156..4211
/organism="SCNN1A"
/standard_name="ENaC"
/number=11
/usedin=Z92978:SCNN1A_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
exon 4359..4434
/organism="SCNN1A"
/standard_name="ENaC"
/number=12
/usedin=Z92978:SCNN1A_cds
/product="amiloride-sensitive epithelial sodium channel
alpha subunit"
intron 4435..4907
/organism="SCNN1A"
/number=12
exon 4908..6313
/organism="SCNN1A"
/standard_name="ENaC"
/number=13
```

/usedin-z92978:SCNNIA_cds
/product="aniloride-sensitive epithelial sodium channel
alpha subunit"

BASE COUNT 1326 a 1778 c 1639 g 1570 t
ORIGIN

Query Match 1.3%; Score 39; DB 10; Length 6313;
Best Local Similarity 100.0%; Pred. No. 3.1e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcctgtccaaactggcagccagtggagcctgtgtgtt 1838
|||||
Db 4916 CCTCCTGTCCAACTGGCAGCCAGTGAGCCTGTGTGT 4954

RESULT 43

HS514K20

LOCUS

DEFINITION

HS514K20 100368 bp DNA PRI 23-NOV-1999

Human DNA sequence from PAC 514K20 on chromosome 6p22.3-24.3. EST,

CA repeats and STS.

ACCESSION

AL008731

VERSION

AL008731.1

KEYWORDS

6p22.3-24.3; repeat polymorphism.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominidae; Homo.

WILDS.

1 (bases 1 to 100368)

Direct Submission

Submitted (09-JAN-1998) Chromosome 6 Project Group

(http://www.sanger.ac.uk/HGP/Chr6/) Sanger Centre, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk

On Jan 13, 1998 this sequence version replaced gi:2598467.

IMPORTANT: This sequence is the entire insert of clone 514K20.

During sequence assembly data are compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variations annotated may not be found in the sequence submission

corresponding to the overlapping clone as we submit sequences with

only a small overlap as described above.

This sequence was generated from part of bacterial clone contigs of

human chromosome 6, constructed by the Sanger Centre chromosome 6

mapping group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6/

This sequence has been finished according to sequence map criteria

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

The true left end of clone 514K20 is at 1 in this sequence. The

true right end of clone 514K20 is at 100368.

514K20 is from the library RPCI3 constructed at the Roswell Park

Cancer Institute by the group of Pieter de Jong.

For further details see <http://bacpac.med.buffalo.edu/>.

Location/Qualifiers

1..100368

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="6"

/map="6p22.3-24.3"

/clone="RP3-514K20"

/clone_lib="RPCI-3"

239..534

repeat_region

/note="AluJo repeat: matches 1..294 of consensus"

1723..2047

misc_feature

/note="match: 267338 STS containing (CA) repeat"

1855..1877

repeat_region

/note="11 copies of CA 100% conserved; differs from

267338"

COMMENT

repeat_region 1882..2174
/note="AluJo repeat: matches 2..301 of consensus"
3503..3808
/note="AluX repeat: matches 2..301 of consensus"
4742..5056
/note="AluJo repeat: matches 1..301 of consensus"
5276..5897
/note="LIMB3 repeat: matches 291..923 of consensus"
6167..6299
/note="LIME3 repeat: matches 767..908 of consensus"
9707..9975
/note="MLT1B repeat: matches 390..113 of consensus"
10357..10552
/note="MLT1C repeat: matches 466..268 of consensus"
10571..11046
/note="MER4A2 repeat: matches 1..505 of consensus"
11042..11286
/note="MLT1C repeat: matches 247..1 of consensus"
12252..12370
/note="MLT1B repeat: matches 5..120 of consensus"
14219..14506
/note="match: 216457 STS containing (CA) repeat"
14279..14319
/note="20 copies of CA 100% conserved; differs from
216457: D6S259"
15184..15235
/note="2 copies of 26 mer 98% conserved"
18638..18929
/note="AluX repeat: matches 302..6 of consensus"
19749..20135
/note="MLT1C repeat: matches 23..466 of consensus"
21690..21761
/note="MIR repeat: matches 186..256 of consensus"
23026..23327
/note="AluX repeat: matches 1..302 of consensus"
23328..23478
/note="AluJb repeat: matches 122..273 of consensus;
incomplete repeat"
24403..24786
/note="LIMD2 repeat: matches 469..884 of consensus"
24779..24937
/note="MER4A2B repeat: matches 1149..1300 of consensus"
26000..26282
/note="AluJb repeat: matches 293..1 of consensus"
29619..29921
/note="AluJ repeat: matches 2..301 of consensus"
31380..31670
/note="AluSg repeat: matches 303..1 of consensus"
31687..31753
/note="MIR2 repeat: matches 141..75 of consensus"
31926..32455
/note="LIMB8 repeat: matches 1035..514 of consensus"
32475..32737
/note="LIMB1 repeat: matches 566..302 of consensus"
32859..33179
/note="AluJo repeat: matches 302..1 of consensus"
33831..33913
/note="MLT1C repeat: matches 89..1 of consensus"
34014..34307
/note="AluJ repeat: matches 1..298 of consensus"
35042..35117
/note="MIR2 repeat: matches 146..74 of consensus"
35186..35487
/note="AluSg repeat: matches 1..303 of consensus"
36019..36447
/note="MLT1C repeat: matches 466..1 of consensus"
38175..38477
/note="AluX repeat: matches 302..1 of consensus"
38520..38643
/note="MIR2 repeat: matches 145..16 of consensus"
39015..39237
/note="LIME1 repeat: matches 432..661 of consensus"
39457..39631

```
/note="AluJo repeat: matches 125. .299 of consensus;
incomplete repeat"
3950. .39741
/note="L1MA2 repeat: matches 964. .1055 of consensus"
40095. .40319
/note="L1PA14 repeat: matches 672. .896 of consensus"
41998. .42078
/note="MIR repeat: matches 144. .56 of consensus"
42369. .42595
/note="MER28 repeat: matches 1. .230 of consensus"
42654. .42911
/note="AluSp repeat: matches 258. .1 of consensus;
incomplete repeat"
42914. .43132
/note="TIGER2 repeat: matches 2472. .2708 of consensus"
43331. .43631
/note="AluTo repeat: matches 1. .300 of consensus"
43633. .43673
/note="MER5A repeat: matches 69. .109 of consensus"
43866. .44166
/note="AluX repeat: matches 1. .301 of consensus"
45157. .45192
/note="18 copies of 2 mer 89 % conserved"
46610. .46915
/note="AluX repeat: matches 302. .1 of consensus"
47459. .47740
/note="AluX repeat: matches 15. .297 of consensus"
48084. .48229
/note="MLT1A2 repeat: matches 1. .149 of consensus"
48230. .48585
/note="MLT1A1 repeat: matches 1. .365 of consensus"
48586. .48820
/note="MLT1B repeat: matches 154. .374 of consensus"
49766. .49791
/note="13 copies of 2 mer 100 % conserved"
49951. .49982
/note="16 copies of 2 mer 88 % conserved"
50250. .50551
/note="AluSp repeat: matches 1. .303 of consensus"
51359. .51922
/note="L1ME3A repeat: matches 630. .58 of consensus"
51805. .51923
/note="L1MA1 repeat: matches 178. .57 of consensus"
52531. .52831
/note="AluSp repeat: matches 3. .302 of consensus"
53351. .53654
/note="AluSp repeat: matches 1. .303 of consensus"
54460. .54843
/note="MSTD repeat: matches 394. .1 of consensus"
55614. .55914
/note="AluSp repeat: matches 1. .303 of consensus"
58484. .58593
/note="MSTC repeat: matches 1. .110 of consensus"
58597. .58702
/note="MSTC repeat: matches 299. .405 of consensus"
59438. .59525
/note="L1MB7 repeat: matches 298. .386 of consensus"
61103. .61222
/note="AluY repeat: matches 1. .124 of consensus;
incomplete repeat"
61225. .62555
/note="PTR5 repeat: matches 2438. .807 of consensus"
62556. .62807
/note="AluY repeat: matches 44. .301 of consensus;
incomplete repeat"
62868. .63041
/note="MER5A repeat: matches 182. .2 of consensus"
64497. .64784
/note="AluJb repeat: matches 297. .1 of consensus"
66089. .66144
/note="MIR repeat: matches 86. .141 of consensus"
67000. .67303
```

```
Query Match 1.3%; Score 39; DB 10; Length 100368;
Best Local Similarity 100.0%; Pred. No. 4.1e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 67277 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 67315
|||||

RESULT 44
AC005840 140026 bp DNA HTG 31-OCT-1998
LOCUS Homo sapiens chromosome 12; between D12S356 and D12S1623 clone
DEFINITION RPC11-102E24, *** SEQUENCING IN PROGRESS ***, 1 ordered pieces.
ACCESSION AC005840
VERSION AC005840.1 GI:37799004
KEYWORDS HTG; HTGS_PHASE2.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 140026)
AUTHORS Montgomery, K.T., Lau, S.T. and Kucherlapati, R.
TITLE High Throughput Sequencing of Human Chromosome 12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 140026)
AUTHORS Montgomery, K.T., Lau, S.T. and Kucherlapati, R.
TITLE Direct Submission
JOURNAL
COMMENT * NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 140026: contig of 140026 bp in length.
FEATURES
source
location/Qualifiers
1..140026
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RPC11-102E24"
/chromosome="12; between D12S356 and D12S1623"
BASE COUNT 35328 a 33235 c 34250 g 37213 t
ORIGIN

Query Match 1.3%; Score 39; DB 41; Length 140026;
Best Local Similarity 100.0%; Pred. No. 4.2e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctctgtccacctgggcagccagctggagcctgtggtt 1838
|||||
Db 136498 CCTCTGTCCACCTGGGCGAGCCAGTGGAGCCTGTGGTT 136536
|||||

RESULT 45
AL158203 157227 bp DNA HTG 20-APR-2000
LOCUS Homo sapiens chromosome X clone RP11-93B10 map q12-13.1, ***
DEFINITION SEQUENCING IN PROGRESS ***, in unordered pieces.
ACCESSION AL158203
VERSION AL158203.2 GI:7634422
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 157227)
```

AUTHORS
TITLE
JOURNAL

Pavitt,R.
Direct Submission
Submitted (19-APR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerquest@sanger.ac.uk
On Apr 21, 2000 this sequence version replaced gi:7160601.
----- Genome Center

COMMENT

Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba93B10
----- Summary Statistics
Sequencing vector: pGAP4; version 4.5
Chemistry: Dye-terminator ET-amersham; 14% of reads Chemistry:
Dye-terminator Big Dye; 85% of reads
Consensus quality: 140176 bases at least Q40
Consensus quality: 146899 bases at least Q30
Consensus quality: 151390 bases at least Q20
Insert size: 154627; sum-of-contigs
Insert size: 125380; 18.5% error; agarose-fp
Quality coverage: 2.97x in Q20 bases; sum-of-contigs Quality
coverage: 3.67x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently * consists
of 27 contigs. The true order of the pieces is * not known and
their order in this sequence record is * arbitrary. Where the
contigs adjacent to the vector can * be identified, they are the
labelled with 'clone.end' in the * feature table. Some order and
orientation information * can tentatively be deduced from paired
sequencing reads * which have been identified to span the gap
between two * contigs. These are labelled as part of the same *
'fragment_chain', and the order and relative orientation * of the
pieces within a fragment_chain is reflected in * this file. Gaps
between the contigs are represented as * runs of N, but the exact
sizes of the gaps are unknown. * This record will be updated with
the finished sequence as * soon as it is available and the
accession number will be * preserved.
*
1 9746 contig of 9746 bp in length; fragment_chain 1 *
9847 29417 contig of 19571 bp in length; fragment_chain 1 *
29518 37118 contig of 7601 bp in length; fragment_chain 1 *
37219 38248 contig of 1030 bp in length; fragment_chain 2 *
38349 48422 contig of 10074 bp in length; fragment_chain 2 *
48523 62333 contig of 13811 bp in length; fragment_chain 2 *
62434 59691 contig of 7258 bp in length; fragment_chain 2 *
69792 71246 contig of 1455 bp in length; fragment_chain 2 *
71347 73427 contig of 2081 bp in length; fragment_chain 2 *
73528 76614 contig of 3087 bp in length; fragment_chain 3 *
76715 80354 contig of 3640 bp in length; fragment_chain 3 *
80455 82645 contig of 2191 bp in length; fragment_chain 3 *
82746 91544 contig of 8799 bp in length; fragment_chain 3 *
93404 contig of 1760 bp in length; fragment_chain 3 *
93505 95394 contig of 1890 bp in length; fragment_chain 4 *
95495 98525 contig of 3031 bp in length; fragment_chain 4 *
98626 104361 contig of 5736 bp in length; fragment_chain 4 *
104462 109443 contig of 4982 bp in length; fragment_chain 4 *
109544 125425 contig of 15582 bp in length; fragment_chain 5 *
125226 127995 contig of 2770 bp in length; fragment_chain 5 *
128096 130614 contig of 2519 bp in length;
* 130715 133455 contig of 2741 bp in length
* 133556 137090 contig of 3535 bp in length
* 137191 138724 contig of 1534 bp in length
* 138825 150575 contig of 11751 bp in length; fragment_chain 6 *
* 150676 152322 contig of 1647 bp in length; fragment_chain 6 *
152423 157227 contig of 4805 bp in length; fragment_chain 6.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
Location/Qualifiers
1. .157227

FEATURES
source

misc_feature
1. .9746
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q12-13.1"
/clone="RP11-93B10"
/clone_lib="RPC1-11.1"
/note="assembly_fragment:00809
fragment_chain:1
clone_end:r7
vector_side:left"
9847. .29417
/note="assembly_fragment:00615
fragment_chain:1"
29518. 37118
/note="assembly_fragment:00573
fragment_chain:1"
37219. 38248
/note="assembly_fragment:00774
fragment_chain:2"
38349. 48422
/note="assembly_fragment:00644
fragment_chain:2"
48523. 62333
/note="assembly_fragment:00969
fragment_chain:2"
62434. 69691
/note="assembly_fragment:00900
fragment_chain:2"
69792. 71246
/note="assembly_fragment:00672
fragment_chain:2"
71347. 73427
/note="assembly_fragment:00184
fragment_chain:2"
73528. 76614
/note="assembly_fragment:00955
fragment_chain:3"
76715. 80354
/note="assembly_fragment:00440
fragment_chain:3"
80455. 82645
/note="assembly_fragment:00594
fragment_chain:3"
82746. 91544
/note="assembly_fragment:00816
fragment_chain:3"
91645. 93404
/note="assembly_fragment:00707
fragment_chain:3"
93505. 95394
/note="assembly_fragment:00151
fragment_chain:4"
95495. 98525
/note="assembly_fragment:00098
fragment_chain:4"
98626. 104361
/note="assembly_fragment:00365
fragment_chain:4"
104462. 109443
/note="assembly_fragment:00981
fragment_chain:4"
109544. 125125
/note="assembly_fragment:00602
fragment_chain:5"
125226. 127995
/note="assembly_fragment:00818
fragment_chain:5"
128096. 130614
/note="assembly_fragment:00295"
130715. 133455
/note="assembly_fragment:00665"
133556. 137090

```
misc_feature /note="assembly_fragment:00849"
137191. .138724
misc_feature /note="assembly_fragment:01034"
138825. .150575
/note="assembly_fragment:00274
fragment_chain:6"
misc_feature /note="assembly_fragment:00558"
150676. .152322
fragment_chain:6"
misc_feature /note="assembly_fragment:00317"
152423. .157227
fragment_chain:6
clone_end:SP6
vector_side:right"
BASE COUNT 49599 a 30099 c 29258 g 45662 t 2609 others
ORIGIN
```

```
Query Match 1.3%; Score 39; DB 40; Length 157227;
Best Local Similarity 100.0%; Pred. No. 4.3e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3115
|||||
Db 97866 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 97904
```

Search completed: September 11, 2000, 01:02:03
Job time: 11111 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:26:52 ; Search time 2749.07 Seconds
(without alignments)
2023.556 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcgaccacgcgtcg.....aaaaaaaaaaaaaaaa 3117

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 892348106 residues
Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_om:*
- 4: gb_ov:*
- 5: gb_pat:*
- 6: gb_ph:*
- 7: gb_pl1:*
- 8: gb_pl2:*
- 9: gb_pl3:*
- 10: gb_pr2:*
- 11: gb_pr3:*
- 12: gb_ro:*
- 13: gb_sts:*
- 14: gb_sy:*
- 15: gb_un:*
- 16: em_fun:*
- 17: em_humi:*
- 18: em_hum2:*
- 19: em_in:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_sy:*
- 29: em_un:*
- 30: em_v1:*
- 31: gb_htg1:*
- 32: gb_htg2:*
- 33: gb_in1:*
- 34: gb_in2:*
- 35: em_ba1:*
- 36: em_ba2:*
- 37: em_hum3:*
- 38: em_hum4:*
- 39: gb_pr4:*
- 40: gb_htg3:*
- 41: gb_htg4:*
- 42: gb_htg5:*
- 43: gb_htg6:*

- 44: gb_htg7:*
- 45: em_htg1:*
- 46: em_htg2:*
- 47: em_htg3:*
- 48: em_hum5:*
- 49: gb_pl3:*
- 50: gb_pr5:*
- 51: gb_htg8:*
- 52: gb_htg9:*
- 53: gb_htg10:*
- 54: gb_htg11:*
- 55: gb_htg12:*
- 56: gb_htg13:*
- 57: gb_htg14:*
- 58: gb_in3:*
- 59: gb_htg15:*
- 60: gb_htg16:*
- 61: gb_htg17:*
- 62: em_htg4:*
- 63: em_htg5:*
- 64: em_htg6:*
- 65: em_htg7:*
- 66: em_hum6:*
- 67: gb_htg18:*
- 68: gb_htg19:*
- 69: gb_htg20:*
- 70: gb_htg21:*
- 71: gb_htg22:*
- 72: gb_htg23:*
- 73: gb_htg24:*
- 74: gb_htg25:*
- 75: gb_htg26:*
- 76: gb_htg27:*
- 77: gb_htg28:*
- 78: gb_htg29:*
- 79: gb_htg30:*
- 80: gb_htg31:*
- 81: gb_vil:*
- 82: gb_vil2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3117	100.0	3117	12	RNASNA	X70497 R. norvegic
2	3062.8	98.3	3081	12	RNASNA	X70521 R. norvegic
3	2280	73.1	3000	12	AF112185	AF112185 Mus muscu
4	2196.8	70.5	2200	12	RNU54700	U54700 Rattus norv
5	2195.2	70.4	2200	12	RNU54699	U54699 Rattus norv
6	1532.2	49.2	3151	9	HSLASNA	X76180 H.sapiens m
7	1436.8	46.1	2334	9	HUMSODIUM	L29007 Human Kidne
8	1401.8	45.0	3586	3	BTU14944	U14944 Bos taurus
9	1375	44.1	2695	12	CP0249296	AJ249296 Cavia por
10	1362.2	43.7	2915	3	OCU132108	AJ132108 Oryctolag
11	1357	43.5	3385	12	AF071230	AF071230 Cavia cob
12	752.2	24.1	3368	4	GGU58475	U58475 Gallus gall
13	751.8	24.1	1003	3	AF232715	AF232715 Ovis arie
14	747.4	24.0	2321	4	GGU62902	U62902 Gallus gall
15	747.4	24.0	3399	4	GGU62903	U62903 Gallus gall
16	729.8	23.4	1065	3	AF229025	AF229025 Oryctolag
17	584.6	18.8	2773	4	XL023535	U23535 Xenopus lae
18	502	16.1	3841	4	GGU62904	U62904 Gallus gall
19	318.4	10.2	1202	10	HSZ92978	Z92978 H.sapiens S
20	318.4	10.2	4526	39	HSANAC1	AF060910 Homo sapi
21	318.4	10.2	140026	41	AC005840	AC005840 Homo sapi
22	313.4	10.1	188488	39	AC006057	AC006057 Homo sapi
23	307.6	9.9	1917	11	HSU38254	U38254 Human amilo
24	305.8	9.8	1917	11	AF038165	AF038165 Pan trogl

25 303.4 9.7 6313 10 HS292981 292981 H. sapiens S
26 303 9.7 2562 39 HSAENAC4 AF060913 Homo sapi
27 295.6 9.5 2448 3 OCUI132110 AJ132110 Oryctolag
28 293 9.4 344 12 MMU52006 U52006 Mus musculu
29 290.6 9.3 2136 12 RNU37539 U37539 Rattus norv
30 290.6 9.3 3012 12 RNEENACG X77933 R. norvegicu
31 288.4 9.3 1923 10 HUMBETA L36593 Homo sapien
32 287.4 9.2 2136 12 RNU37540 U37540 Rattus norv
33 287.4 9.2 2938 12 RNNAC X78034 R. norvegicu
34 286.8 9.2 2564 9 HSCNN1B X87159 H. sapiens m
35 284 9.1 2436 12 AF112186 AF112186 Mus muscu
36 283.6 9.1 749 12 AF082073 AF082073 Rattus no
37 283.6 9.1 2223 12 AF081783 AF081783 Rattus no
38 282.8 9.1 2989 12 AF112187 AF112187 Mus muscu
39 278.4 8.9 2221 12 RNU35174 U35174 Rattus norv
40 276.8 8.9 2221 12 RNU35175 U35175 Rattus norv
41 276.8 8.9 2462 12 RNEENACB X77932 R. norvegicu
42 274.6 8.8 2204 10 HUMGAMM L36592 Homo sapien
43 272.6 8.7 2495 3 OCUI132109 AJ132109 Oryctolag
44 271.4 8.7 3384 9 HSCNN1G X87160 H. sapiens m
45 262.2 8.4 3022 4 XLU25342 U25342 Xenopus lae

ALIGNMENTS

RESULT 1
RNEACA RNEACA 3117 bp mRNA ROD 08-MAR-1994
R. norvegicus mRNA for sodium channel, alpha subunit.
X70497
X70497.1 GI:458845
alpha subunit; sodium channel.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 3117)
Canessa, C.M., Horisberger, J.D. and Rossier, B.C.
Epithelial sodium channel related to proteins involved in
neurodegeneration
Nature 361 (6411), 467-470 (1993)
93156815
Canessa, C.M.
Direct Submission
Submitted (02-MAR-1994) C.M. Canessa, Institut de Pharmacologie et
Toxicologie, Bugnon 27, 1005 Lausanne, SWITZERLAND
Location/Qualifiers
1. 3117
/organism="Rattus norvegicus"
/strain="Sprague Dawley"
/db_xref="taxon:10116"
/dev_stage="adult"
/tissue_type="distal colon epithelium"
82. 2178
/gene="rEnac"
82. 2178
/gene="rEnac"
/codon_start=1
/product="epithelial sodium channel, alpha subunit"
/protein_id="CAA49905.1"
/db_xref="GI:458846"
/db_xref="SWISS-PROT:P37089"
/translation="MLDHTRAPELMIDLHASNPKSMKMQKEQDPPQPMQG
LGKDKREQGLGPFSPAPQPTFEEALIEPHRSYRELQFCNNTIHGAIRLVCS
KINRMTAFWMLCTFCGMVWQFALLFEYLYSPVSLINLSNDKLVFPVAVTCL
NRYRTEKEELEELDRTEOTLFDLYKNSSYTRQAGARRSSRDLLGAFPHPLQRL
RPPPPYSRTARSSSSVRDNNPOVDKDKWIGFQLCNKSDCFYOTYSSVDVAVR
EWRHYINLSRSDTSPLAEEALGNFIFCRFNQAPCNQANYSKHHPMYGNCT
FNDKNNLSWMSFVNGNLSLTIRTEQNDIFPLSTVTGARVWVHGDEPAFMDGD
GFNLRPVTSSIMRKEALDSLGNYGDCETENGSDVPVKNLYPSKYTQQVCVTHSCFQE
NMIMKGCAYIFYPKPGVEFYDRKQSSWGYCYKLOGAFSLDSLGCFSKRKPFCV

FEATURES

source
gene
CDS

BASE COUNT 684 a 983 c 753 g 697 t
ORIGIN
Query Match 100.0%; Score 3117; DB 12; Length 3117;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cgggtgcaccacgcgtcgcgagggcagcctggatgcggcgacggtccccgagccccc 60
|||||
Db 1 CGGGTCGACCCACCGTCGCGAGGCGACGCTGGATGGCGGACGCGTCCCGAGACGCC 60
QY 61 ctttgccttcacgcctaatgatgtgtgacacacacacagacccctgagctcaacattgac 120
|||||
Db 61 CATTCTGCTTCACGCTAATGATGTGGACACACACAGAGCCCTGAGCTCAACATTGAC 120
QY 121 cttagaccttcacgcctccaactcgcctaaggggtccatgaagggaaccaattcaaggag 180
|||||
Db 121 CTAGACCTTCACGCTCCAACTCGCTAAGGGGTCCATGAAGGGCAACCAATTCAAGGAG 180
QY 181 caagaccttgcctcctcctcagcccatgcaagactggggagggggacaaaagtggaagag 240
|||||
Db 181 CAAGACCTTGTCTCTCTCAGCCCATGCAAGACTGGGGAAGGGGACAAAGTGAAGAG 240
QY 241 caggccctggggcccggaacccctcagcaccccgggcagccacacgagagagagagcactg 300
|||||
Db 241 CAGGGCTTGGGCGCGGAAACCTCTCAGCACCCCGGACCCACCGAGGAGGAGGACACTG 300
QY 301 attgaattccacgcctcctaccggagctcttccagttcttctgcaacacacaccaccatc 360
|||||
Db 301 ATGAATTCACCGCTCTCTACCGGGAGCTCTTCCAGTTCTTCTGCAACCAACACCACATC 360
QY 361 caggggccatccgcctgggtgtgtctccaaacacacacacacacacacacacacacacacac 420
|||||
Db 361 CAGGGGGCCATCCGCCCTGGTGTCTCCAAACACAAACCGCATGAAGACGGCTTCTGGGGG 420
QY 421 gtgtgtggtgtgacaccttcggcagatgactgagtcagtcagtcagtcagtcagtcagtcag 480
|||||
Db 421 GTGCTGTGGCTGTGCACCTTCGGCATGTACTTGGCAGTTGCTTGTGTGGAGGAG 480
QY 481 taacctcagctaccagtgagctcacaactcaactcaactcaactcaactcaactcaactccct 540
|||||
Db 481 TACCTCAGTACCTACCTGAGCTCAACATCAACCTCAATTCAGACAAAGCTGGTTCCTTCC 540
QY 541 gcgctcactgtctgcacccttaactccttacagatacactgaaattaaagagagagctggaa 600
|||||
Db 541 GCGGTCACTGTCTGCACCCCTTAATCCTTACAGATACACTGAAATTAAGAGAGAGCTGGAA 600
QY 601 gagctgacgcgatacagcagagacgcttttgaactgtacaaaatacactcttctctac 660
|||||
Db 601 GAGCTGGACCGCATCAGGAGCAGAGCGCTTTTGTGACTTGTACAAATACAACTTCTCTAC 660
QY 661 actcgcagctggggcccgagcccgagctcccgacccctcctcctcctcctcctcctcctc 720
|||||
Db 661 ACTCGCCAGGCTGGGGCCCGACGCCGAGCTCCCGGACCTCCTCGGTGCTTTTCCGCCAC 720
QY 721 ccctgcagcgctgcgcacccctccactccgcctcctcctcctcctcctcctcctcctcctc 780
|||||
Db 721 CCGCTGCAGCGCTTGGCAGCTTCCACCTCGCCCTTACTCGGGCGCAGCGCGCAGCGGG 780
QY 781 tcttcacgcgtacgcgac 840
|||||
Db 781 TCTTCCAGCGTACGCGACACAAATCCCAAGTGACCGGAGGAGACTGGAAGATTCGGCTTC 840
QY 841 caactgtcaaccagac 900
|||||
Db 841 CAACGTGTCACACACACAAATCAGACTGTTTCTACAGACATACATCTCTCTCGGGTGGAT 900
QY 901 gcagtgaggagtggtaccgcttccattacatcaacattctgtccagactgtcggacacc 960

|||||
Db 901 GCAGTGGAGAGTGTACGGCTTCCATTACATCAACATTCTGTCCAGACTGTGCGACACC 960
QY 961 tcgcccgtctagagaaagaagccctgggcaacttctcatcttcaacctgtcgcttcaaccag 1020
Db 961 TCGCCCGCTCTAGAGAAAGAGCCCTGGGCAACTTCATCTTCACTTCCCTGGCTTCAACCCAG 1020
QY 1021 gccctctgaacagcgcaattattccaagttccacacccccaatgtacgggaactgctac 1080
Db 1021 GCCCCCTGCAACACAGCGCAATATTCCAAGTTCCACACCCCCATGTACGGGAACCTGCTAC 1080
QY 1081 actttcaatgaacaagaactccaatctcttgatgtcctccaatgacctggagttcaacaat 1140
Db 1081 ACTTTCAATGACAAGAACAACCTCAATCTCTGGATGTCCTCCATGCTGCGTGGAGTCAACAAT 1140
QY 1141 gatttgcctcacactgcacagacagagaatgacttcaatccccctgtgtccacagt 1200
Db 1141 GGTTTGTCCCTGACACTGCGCACAGAGCAGATGATCTTATCCCTTGTGTCCACAGTG 1200
QY 1201 acggggccaggggtgatggatggtcaggatgagccctgcctttatggatgatggc 1260
Db 1201 ACGGGGCCAGGGTATGATGATGTGTAGGATGAGCCTGCCTTTATGGATGATGTGGC 1260
QY 1261 tbaacttgagccctggcgtgagacccccaatccatccatgatatgagaaagaagccctggacagc 1320
Db 1261 TTCAACTTGAGGCTGGCTGGAGACCTCCATCAGTATGAGAAAGGAAGCCCTGGACAGC 1320
QY 1321 ctccgagaaataacagcgactgtactgagaatggtagcgaagtcccggtccaagaacctt 1380
Db 1321 CTCGGAGGAAATTTACGGCGACTGTACTGAGAAATGGTAGCGATGTCCCGGTCAAGAACCCTT 1380
QY 1381 tacccttccaagtatacacagcaggtgtgcatctcactcctgctccagagacaatgatc 1440
Db 1381 TACCCTTTCCNAGTATACACAGCAGGTGTGCATTCATCTCTGTTCCAGGANAACATGATC 1440
QY 1441 aagaagtgtgctgtgcctacatcttctaccctaaagcccaaggagttgagttctgtgac 1500
Db 1441 AAGAAGTGTGGCTGTGCCCTACATCTTCTACCTTAAGCCCAAGGAGTTGAGTTCTGTGAC 1500
QY 1501 taccgaagacagagctccctggggctattgctattataaactgcagggcccttctccttg 1560
Db 1501 TACCGAAACAGAGCTCCCTGGGGCTATTGCTATTATAAATGCAGGGCGCCTTCTCCTTG 1560
QY 1561 gcagcctggcgtgttctccaagtcggaagcccttgtagtgtatcaactacaactc 1620
Db 1561 GACAGCCTGGGCTGTTTCTCCAAGTGTCCGAGCCCTTGTAGTGTGATCACTACAACATC 1620
QY 1621 tctgcggctactcacgggtggccatctgtgaagtccaggattggatcttccagatgctg 1680
Db 1621 TCTGCCGGCTACTCACGGGTGGCCATCTGTGAAGTCCCAGGATTGGATCTTCGAGATGCTG 1680
QY 1681 tccctgcagaacaataactattacaacaagaagaacagggagttgcgaagctccaacatc 1740
Db 1681 TCTTTGCAGAACAATTTACACTATTACAACAAGAAGAACGGAGTTGCAAAAGCTCAACATC 1740
QY 1741 ttctcaaggagctaaactataaaactaattcggagctcctctctgtccagatggtcagc 1800
Db 1741 TTCTTCAAGAGCTGAACATATAAACTAATTCGGATGTCTTCTGTCTACGATGGTCAAG 1800
QY 1801 ctctgtccaacctgggagccagtgaggcctgtggttttggctcgctcgctcctctgtg 1860
Db 1801 CTCTGCTCAACCTGGGACAGCAGTGGAGCCGTGTGTTGGCTCGTCCGCTCTCTGTG 1860
QY 1861 gtygagatggcgagctcatcttcacacctcctgggtccatcacacttctatgtcgtacgc 1920
Db 1861 GTGGAGATGGGACGCTCATCTTCGACCTTCTGGTGTGTTGGCTCGTCCGCTCTCTGTG 1920
QY 1921 cggttccggagcgggtactggtctccaggagcagggggccaggggtgccaagggaggtggcc 1980
Db 1921 CGGTTCCGGAGCGGTACTTGGTCTCCAGACAGGGGGCCAGGGGTGCCAGGGAGTGGCC 1980
QY 1981 tccactccaggtctctcctcccgctcccgcttctgtccctcacacctacatccccaccact 2040
|||||

Db 1981 TCCACTCCAGCTTCTCTCCGCTCCCGTTTCTGTCTCTCACCTACATCCCCACCACCT 2040
QY 2041 tctttcccccagagggcatgacccctccctggcctgagccctcacagccctcacagctcat 2100
Db 2041 TCCTTGGCCCCAGAGGGCATGACCCCTGCCCTTCCCTTGACAGCCCTCCACCTTGCCTAT 2100
QY 2101 gctactctaggccccagtgccccctccactggaactctgctggcgcgtgactgttctgctgt 2160
Db 2101 GCTACTCTAGGCCCCAGTGCCCTCCACTTGGACTCTCGCGGCGCTGACTGTCTCGCCTG 2160
QY 2161 gccctggcgcgtctgagagagagaagatcctctcaaccagccctgagctccttgt 2220
Db 2161 GCCCTGGCGCGCTCTGAGAGAGAGAAGGATCTCTACGCCAGCCCTGAGCTCCCTTGT 2220
QY 2221 aaactatgatatctcaacttcagcacatcttcccagcggcccccgagccctctgtg 2280
Db 2221 AAATGATGATATCTACTTTCAGCATCTTCCCCAGCGCCCCCGGAGCCCTCTGTG 2280
QY 2281 gttgctgcctgaggaacaggtctgtagtaaggggccaggaagcgtccagaggacag 2340
Db 2281 GTGTGCTGCTGAGGGAACAGGCTGAGTAAGGGGCCAGGAAGCCGTCCAGAGGACAGG 2340
QY 2341 ggttaacgatctgctcagagcgtccctgccccctgtctctgaacactgccttccaagca 2400
Db 2341 GGCTAAGATCTGCTCAGAGCTGCCCTGCCCTGCTCTTCTGAACACTGCCTTCCACAAGCA 2400
QY 2401 cagacaagttcccttccctcgatcagcccagccagacttgagctcgacaagaac 2460
Db 2401 CAGACAAGTTCCCTCTTCCCTCGATCAGCCCCAGCCAGACTTGGAGCTCTGACAAGGAAC 2460
QY 2461 ttccctgggagaggaaccaaggaacaaatagaacagccatcgactcctcaacccgtgacc 2520
Db 2461 TTTCTTGGGAGAGGACCAAGGAACAATAGAAGCCATCGCACTCTTACCCCGCTGACC 2520
QY 2521 agaggtggcctggcctcactgccttcagggacactgagtctgtcctcctccttgaact 2580
Db 2521 AGAGGCTGGCTGGCCTCACTGCCTTTCAGGGACACTGATGCTGCTCTCTCTTGAAT 2580
QY 2581 tgggtggggaaccccaacaaagccctttagttcttggcaattccctccctc 2640
Db 2581 TGGGTGGGGAACCCCAACCCAAAGCCCTTGTAGTTCCTTGGCAATCCCCCTCCCTC 2640
QY 2641 actcctcaggtgggactagagtaagacagaccaggttaggctgtgtctctctcat 2700
Db 2641 ACTCCTCAGGCTGGGACTAGAGTAAACAGACAGACCAGGCTAGGCTGTCTCTCAT 2700
QY 2701 tctcaagactctctcatcttgatacacatccccagtcgctctgtgtgtgtctctc 2760
Db 2701 TCTCAAGACTCTCTCATTTGATACCATACCCAGTGCCTCTGTGTGTGTCTCTCTTC 2760
QY 2761 tttctgttctcagcctggaagcttcttcaactgttccctcagaacaacagactctatgca 2820
Db 2761 TTTCTGTCTCTCAGCCTTGGAGCTTCTTTCACTGTCCCTCAGAACACAGCTCTATGCA 2820
QY 2821 tcccttagaaacctgccaaagccacattcttctgtgaactcagttaccctgtatgt 2880
Db 2821 TCCCTTAGAAACCTGCCAAAGGCACATTCTTTGTGTGAACTCAGTTACCCCTGTATGT 2880
QY 2881 ctccccagaattgtgcgcccctctccccctacccccacagtgactgtattatattgttccac 2940
Db 2881 CTCCCCAGAAATGTGCGCCCTCTCCCTCACCCCCACAGTGTACTGTATTATATGTTTAC 2940
QY 2941 attcttctgtctgcctccctgggttaaacctgagctccttctgtgtgcagggatgagatt 3000
Db 2941 ATTCTTTTGTGTGCTGCCCTGCCCTGGTAAACTGAGCTCCTTGTGTGTGTCAGGGATGAGATT 3000
QY 3001 tgcctgttttctatcttccggtctagcccagcttcccacttgggagcgggtgagcaggt 3060
Db 3001 TGTCTGTGTGTATCTTCCGGTCTAGCCAGTCTCCCACTTGGGACGGGTAGGAGGT 3060
QY 3061 actcaataaagggtgtgttccatcaaaaaaataaaaaaataaaaaaataaaaaa 3117
Db 3061 ACTCAATAAAGGCTTGTTCATCAAAAAAATAAAAAAATAAAAAAATAAAAAA 3117

Db 1262 CTCCTCATAGTATGAGAAAGGAAGCCCTGGACAGCCTCGGAGGAAATATACGGCGACTGTAC 1321
Qy 1347 tgagaatggttagcgtgtcccggtcaagaacctttacccttccaagtatacacagcaggt 1406
Db 1322 TGAGAAATGGTAGCGATGCCGGTCAAGAACCTTTACCTTCCAGTATACACAGCAGGT 1381
Qy 1407 gtgcattcaactctcttccaggagaaacatgatcaagaagtgtggtgtgcctacattt 1466
Db 1382 GTGCATTTCACTCCTGCTTCCAGGAGAAATGATCAAGAAGTGTGGCTGTGCCTACATCTT 1441
Qy 1467 ctaccctaagcccaaggaggttgagttctgtgactacgaagacagagcctctggggcta 1526
Db 1442 CTACCTTAAGCCCAAGGGAGTTGAGTTCTGTACTACCGAAAGCAGAGCTCTGGGGCTA 1501
Qy 1527 ttgctattataaactgcaggcgccctctccttgacagcctgggctgtttctccaagtg 1586
Db 1502 TTGCTATTATAAATGACAGGGCGCCTTCTCCTTTGGACAGCCTGGGCTGTTCCTCAAGTG 1561
Qy 1587 tcgggaagcctttagtgtgatacaactcaaaactctctgcggctactcacggtggccatc 1646
Db 1562 TCGGAAGCCTTGTAGTGTGATCAACTACAAACTCTCTGCGGCTACTCACGGTGGCCATC 1621
Qy 1647 tgtgaagtcceagattgagatttcgagatgctgtccttgcagaacaattacactattaa 1706
Db 1622 TGTAAGTCCCAAGGATTCGAGATGCTGCTTTCGAGATGCTCTTTCGAGAACAATTACACTATTAA 1681
Qy 1707 caacaaagaacgaggtgcaagctcaacatcttctcaagagagctgaactataaac 1766
Db 1682 CAAACAAAGAACGGAGTTGCAAACTCAACATCTCTTCAAGGAGCTGAACATATAAAC 1741
Qy 1767 taatcggagtgctctctgtcacgatggtcagcctctgtccaaactgggagcagtg 1826
Db 1742 TAAATCGGAGTCTCCTTGTGTACAGATGGTCAGCCTCTGTCTCAACCTGGGCGCCAGTG 1801
Qy 1827 gagcctgtggttgctcgtcgtctctgtgtgtgagatggtgcggaagtcattctoga 1886
Db 1802 GAGCCTGTGTTGTGCTGCTGCTCTCTGTGTGTGAGATGGGGAGCTCATCTTCA 1861
Qy 1887 cctcctgtcatcacacttctcatgctctacgctacgctggttccgagaccggtactgctcc 1946
Db 1862 CCTCCTGTGTATCACACTTCTCATGCTGTCTACGCCGGTTCGGAGCGGCTACTGTGCTCC 1921
Qy 1947 aggcagaggggcaggggtgcagggaggtggcctccactccagcttctccttcccgctc 2006
Db 1922 AGGACGAGGGGCCAGGGGTGCCAGGAGGTGSCCTCCACTCCAGCTTCTCCTTCCCGTC 1981
Qy 2007 cgtttctgctcaacctacaatcccaaccactcttcttggcccgagggcaatgacccc 2066
Db 1982 CGGTTTCTGTCTCCTCCTACATCCCACTCCACCTCTTTTGCCCCAGCAGGGCATGACCCC 2041
Qy 2067 tccctggcctgacagccctccacgtcctatgctactctagcccccagtgcccctcc 2126
Db 2042 TCCCTTGCCCTTGACAGCCCTTCCACTGCTATGCTACTCTAGGCCCCAGTGCCCTCC 2101
Qy 2127 actggaactctgcggcgctgaacttctgctgtgctcctggcggtcgtctgagagaggag 2186
Db 2102 ACTGAGCTCTGGCGGCGCTGACTGTCTGCTGTGCTGCTGGCGGCGCTCTGAGAGAGGAG 2161
Qy 2187 aaggatcctctcaaccagccctgagctccctgtgaaactgtatgatatctcaactcaga 2246
Db 2162 AAGGATCTCTCACCGACCTTGAGCTCCTTGTAACTGATGATATCTCATCTTCAGCA 2221
Qy 2247 catcttccccagggccccgcgagccctctgtgtgtcctgctgagggagaaaggctg 2306
Db 2222 CATCTTCCCCAGCGGCCCGCGAGCCCTCTGTGTGTGTGCTGTGCTGAGGGAACAGGCTG 2281
Qy 2307 agtaaggggccaggaagccgtccagagagagggctaaagatctgctcagagctgcc 2366
Db 2282 AGTAAGGGGCCAGGAAGCCGTCCAGAGACAGGGGCTTAACGATCTGTCTAGAGCTGCC 2341
Qy 2367 tgcctcctgtcttgaactgccttc--cacaagcacagacaagttccctcttccctgg 2424
Db 2342 TGCCCTGCTTCTGAACACTGCCTTCCACACAAGCACAGACAAGTTCCTCTTCCCTCGG 2401

Qy 2425 atacccccagcagacttgagctctgacaagaaactttcctgggagagaccacgaagac 2484
Db 2402 ATCAGCCCGACGAGACTTGGAGCTCTGACAGGAACCTTCTCTGGGAGAGACCAAGAAC 2461
Qy 2485 aaatagaacagccatgcgactcctaccctcgtagccagaggtggcctggcctcaactgcc 2544
Db 2462 AAATAGAACAGCCATCGCACTCTACCCCGTGACCAGAGGCTGGCTGGCCTCACTGCC 2521
Qy 2545 ttccagggaacagtgtctgctcctcctctgaacttgggtggggaaccccccaaaaag 2604
Db 2522 TTACGGGACACTGATGTGCTCTCTTGAACCTTGGGTGGGGAACCCCAACAAAAG 2581
Qy 2605 cccctctgttagttctttggcaattccctccctcaactcctcaggtggtgggactagagt 2664
Db 2582 CCCCTTGTGTAGTCTTTTGGCAATTCCTCCCTCCTCCTCAAGACTCTCTCATTTGTATAC 2641
Qy 2665 aagacagaccaggttagcctgggtgtctctctctctcaagactctctcatattgatac 2724
Db 2642 AAGACAGACAGCGTAGGCTGGTGTCTCTCTCATTTCTCAAGACTCTCTCATTTGTATAC 2701
Qy 2725 accataccccagtgctcctgtgtgtgtgtctcctctctctctcagcctggaaagc 2784
Db 2702 ACCATACCCCAAGTGCCTCTGTGTGTGTCTCTCTCTTCTTCTTCTCCTCAGCCTGGAAGC 2761
Qy 2785 ttcttctactgtccctcagaacacacacagctctatgcatcccttagaaccctgccccaaagcc 2844
Db 2762 TTCTTTCACTGTCCCTCAGAACACACAGCTATATGCAATCCTTAGAANACCTGCCCAAGGC 2821
Qy 2845 accattcttttgaactcagttaccctgttatgtctcccccagaaattgtcgccctctc 2904
Db 2822 ACCATTTCTTTGTGAACTCAGTTACCTGTATTATGCTCCCCCAGAAATGTGCGCCCTCTC 2881
Qy 2905 cctcaacccccacagtgactgtattatatgttcaattcttttctgtctgcctccctgg 2964
Db 2882 CCCTCACCCCCACAGTGACTGTATTATATATGTTCACATTCCTTTTGTGTCTGCTCCCTGG 2941
Qy 2965 gtaaacctgagctcctgtgtgtcagggatgagattgtctgtttgtatcctcctcgg 3024
Db 2942 GTAAACTGAGCTCCTGTGTGTGTGTCAGGGATGAGATTTGCTGTGTGTATCCTCCGGT 3001
Qy 3025 ctagccagctctccccacttgggagcggtaggcaggtactcaataaaaggctgttccatca 3084
Db 3002 CTAGCCCACTCTCCCACTTTGGGACGGTAGGCAGGTACTCAATAAAGGCTTGTTCATCA 3061
Qy 3085 aaaaaaiaaaaaaiaaaaaa 3104
Db 3062 AAAAAAIAAAAAIAAAAAIA 3081

RESULT 3

AF112185
LOCUS AF112185 3000 bp mRNA ROD 27-MAR-2000
DEFINITION Mus musculus epithelial sodium channel alpha subunit mRNA, complete cds.
ACCESSION AF112185
VERSION AF112185.1 GI:4469398
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3000)
AUTHORS Ahn,Y.J., Brooker,D.R., Kosari,F., Harte,B.J., Li,J., Mackler,S.A. and Kleyman,T.R.
TITLE Cloning and functional expression of the mouse epithelial sodium channel
JOURNAL Am. J. Physiol. 277 (1), F121-F129 (1999)
MEDLINE 99345762
REFERENCE 2 (bases 1 to 3000)
AUTHORS Ahn,Y.J., Brooker,D.B. and Kleyman,T.R.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1998) Renal Electrolyte Division, University of

Pennsylvania, 422 Curie Boulevard, 700 Clinical Research Building,
Philadelphia, PA 19104, USA

FEATURES

source
Location/Qualifiers
1. :3000
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="6"
/map="distal"
32. :2131

CDS

/codon_start=1
/product="epithelial sodium channel alpha subunit"
/protein_id="AAD21244.1"
/db_xref="GI:4469399"
/translation="MDHTRAPELNLDLDVNSPKSMKGNKFKEDLPPLPMQG
LGKDKREQLGPPSPPEPEEALIEFHSRYELFOFFCNNTTHGAIRLVCS
KHNRMTEAFWLEIDRTQMYWQFALLEFYFYPVSLNINLNSDKLVPFAVVTCL
NPYRTEIKARSALEEDLRTQTEOLFYLKYNSTYRQAGRRSRDLGALPHPLQRL
RTPPPNIPARSASSSVRRNPNQVDRDKMKIGFOLCNONKSDCFYQYSSGVDAVR
EWRPHYINILSRPDTSPALEEALGSFIETCRENOAPCQANYSQRIHPMYGNCYT
FNKNNSNMWSSMPGVNGLSLTURTQNDIFPLSLTVTGARVMWVODEPAFMDGG
GFNVRPGETSISMKEALDLSGGNYGDTENGSDVPKNLYPSKYTOQVCIHSQFE
NMIRKCGAYIFYPKPKGVFECDYLKQSWGYCYKQLQAFSLDSLGSFCKRKPCSV
TNYLSAGYSRWPSVKSDWIFEMLSLQNNYTIINNRNGVAKLNIFFRLNRYKNTSES
PSVTMVSLSNLGWSWLFSSVLSVVEMAELIFDLIVITLIMLHFRSRYWSPGR
GARGABEVASTPASSFPSPFCPHTSPPPSLPQOGTTPPLALTAPPPAYATLGPSPSP
LDSAVPSSACAPANAL"

BASE COUNT 666 a 958 c 701 g 675 t

ORIGIN

Query Match 73.1%; Score 2280; DB 12; Length 3000;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 2674; Conservative 0; Mismatches 300; Indels 104; Gaps 11;

QY 51 cgagacagccattctgcctcaagcgtacgtatgatgctggaccacacagagccctgagct 110
DB 1 CGGACGGCCCCATTCTGCCTCCATCTACTAATGATGCTGGACCACACAGGCCCTCAGCT 60
QY 111 caacttgacctagacctcaagcctcccaactcgcctaaaggggtcctgaagggcaacaa 170
DB 61 CAACCTTGACCTAGACCTTGACCTGCCAACTCACCAGGAGGATCCATGAAGGGCAACAA 120
QY 171 attcaaggagcaagacctgtcctcctcagcccatgcaagactgggaaggggcaacaa 230
DB 121 TTCAAGGAGCAGACCTTTGCTCTCTCTGCCCCATGCAAGGACTGGGCAAGGGGGACAA 180
QY 231 acgtgaagagcaggcgtgggcccgaacccctcagcaccgcccgagccacccagagagga 290
DB 181 GCGTGAAGAACAAGCGCTGGCGCCGGAACCTTCAGAGCCCGCGCACGCCACGAGGAGGA 240
QY 291 ggaggcaactgattgaattccacgcctctacacgggagctcttcagttcttctgcaacaa 350
DB 241 GGAGGCACTGATCGAGTTTCACCGCTCTCACCAGGAGCTTCCAGTTCTTCTGCAACAA 300
QY 351 caccacatccacggggccatcgccgtggtgtctcacaacacacacacacacacacacgc 410
DB 301 TACCACCATCCAGGGTGCATCCGCTGGTGTGCTCCAGACCAACCGCATGAAGACGGC 360
QY 411 cttctggcggtgctggtggtgacccctcgccatgctgactgagcttcgaccttgccttgc 470
DB 361 CTCTCGCGGGTCTGCTGCTCGACCTTCGCGCATGATGTAAGTCTGCTTTCCTTTCCT 420
QY 471 gttcggaggagtaacctcagctaccagctgagcctcaacatacaacctcaattcagacaagct 530
DB 421 GTTCGAGGAATAATTTTACCTACCCCGTGAAGTCTCAACATCAACCTCAATTCGGCAAGCT 480
QY 531 ggtcttccctcgctgactgctgaccccttaactcctcaacatacacacacacacacacaa 590
DB 481 GGTCTTCCCTCGCGTCACTGTGTGACCCCTTAATCTCTACAGATACACTGAAATTAAGA 540
QY 591 ggagctgggaagagctggaccgcatcacgagcagacgcttttgcattgtacaaatacaa 650
DB 1621 CGAGATGCTATCTCTGCGAATAAATTACACGATCAACAAACAAACAAACGAGTTGCTAA 1680

DB 541 GGATCTGGAAGAGAGCTGGACCGCATACGGAACAGACGCTTTTGAACCTGTACAAATACAA 600
QY 651 ctcttctacactcgccaggctggggcccgagcgcgcagctcccgagacctctcgtgggtgc 710
DB 601 CTCTTCTTACTACTCGCAGGCTGGGGCGCGCCGACGACACCGGACCTTCGGGGGTGC 660
QY 711 ttcccgacccccctcagcgctcgccatccacactccgcctactcctcgccgcgcgcgcgc 770
DB 661 TCTCCACACACCCCTGACGGCTCGGCACACACCTCCGCCCAATCCCGCCGCTCGGC 720
QY 771 ggcagcggtgtctccagcgacgcgacacacacacacacacacacacacacacacacacac 830
DB 721 GCGCAGCGGCTCTTCCAGTGTACGCGACACAAATCCCAAGTGCAGAGGAGGACTGGAA 780
QY 831 gatcggtctccaaactgcaaccagacacacacacacacacacacacacacacacacacac 890
DB 781 AATCGGCTTCCAACTGTGAACACAGAACAAATCAGACTGCTTCTACGACACATACTCATC 840
QY 891 tggggtggtgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 950
DB 841 CGGGGTGGATGCGCTGAGAGAGTGGTACCGCTTCCATTTACATCAATTTCTGTCCAGACT 900
QY 951 gtccgacacctcgcccgctcgtagaggaagagccctggggcaactctcatcttccacctgtcg 1010
DB 901 GCCCGACACCTCGCCCTGCTAGAGGAAGAAGCCCTGGGCAGCTTCATCTTTACCTGTGC 960
QY 1011 cttcaacacggggccctcgcaaccagcggaatttctcaagttccacacacacacacacacac 1070
DB 961 TTTCAACACGAGCCCTGCAATCAGGCGAATATTCTCAGTTCACACACCCCTGTATGG 1020
QY 1071 gaactgtacactttcaatgacaagaacacacacacacacacacacacacacacacacacac 1130
DB 1021 GAACCTGTACACTTTCACACACAGAACAACTCCAATCTCTGATGCTTTCATGTCCCTGG 1080
QY 1131 agtcaacaatggttgcctcgacactgcgcagacagacagacacacacacacacacacacac 1190
DB 1081 AGTCAACAATGTTTGTCTCCCTGACACTGCGCACAGAGCAGAACTGACTTATCCCTGCT 1140
QY 1191 gtccacagtgaacggggccaggggtggtggtggtggtggtggtggtggtggtggtggtggtg 1250
DB 1141 GTCACAGTGAACGGGGCCAGGGTGTGTCAGCGGTGTCAGGATGAGCGCTCTTTATGGA 1200
QY 1251 tgatggtggttcaacttgagcgctggtggtggtggtggtggtggtggtggtggtggtggtg 1310
DB 1201 TGATGTTGCTTCAAGTGAAGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1260
QY 1311 cctggacagcctcgaggaataattacgcgagactgtactgagaatggtagcgtatgcccggt 1370
DB 1261 CTTGGACAGCTCGGAGGCAACTACGAGAGACTGCACCTGAGAATGGCAGCGATGTCCCTGT 1320
QY 1371 caagaacctttaccttcaagtatacagaaggggtggtggtggtggtggtggtggtggtggtggt 1430
DB 1321 CAAGAACCCTTACCCCTCCAAAGTACACACAGAGGTTGTCATTCACCTCTCTCTCTCCAGGA 1380
QY 1431 gaacatgatacaagaagtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1490
DB 1381 GAACATGATCAAGAAGTGTGGCTGTGCTACATCTTCTACCTTAAGCCCAAGGGGTGTAGA 1440
QY 1491 gttctgtgactaccgaagcagagctcctgggggtattgtctattataaactcagcaggcgc 1550
DB 1441 GTTCTGTGACTACTTAAAGCAGAGCTCCTGGGGCTACTGTACTATATAAATGCGAGGCTGC 1500
QY 1551 cttctccttggaacagcctgggtgtttctccaaagtgcggaagcctgttagtgtgacaa 1610
DB 1501 CTTTCTCTTGGATAGCTGGCTGGCTGCTTCTCCAAAGTCAGGAAGCCGTCGTCAGTGTGACAA 1560
QY 1611 ctacaactctctcgccgctactcagtggtggtggtggtggtggtggtggtggtggtggtggtg 1670
DB 1561 CTACAGCTCTCTGCTGGCTGCTCAAGATGCGCTGCTGTGAAGTCCCGAGGATTTGGATCTT 1620
QY 1671 cgagatgctgtccttgagaac 1730
DB 1621 CGAGATGCTATCTCTGCGAATAAATTACACGATCAACAAACAAACAAACGAGTTGCTAA 1680

[illegible]

Qy	1731	gctcaaatctcttccaagagctgaactataaactaatctcgagctctcttctgcaac	1790
Db	1681	ACTCAACATCTCTTCAAGAGAGCTGAAGCTATAAACTAAATTCGGAGTCTCCCTCTGTCAAC	1740
Qy	1791	gatggtcagctctctccaaacctggcagccagtcagagctggtgttgctcgctccgt	1850
Db	1741	GATGGTCAGGCTCTGTCTCCAACTTGGGACGCCAGTGGAGCTGTGGTTCGGCTCATCTGT	1800
Qy	1851	gctctctggtgagatggcgagctcatctctcgacctctctgctcatcacactctcat	1910
Db	1801	GCTGTCCGTGGTGAATGGGAGCTCATCTTCGACCTCCTGTGTATCACACTCATCAT	1860
Qy	1911	gctgctacgcoggtctcogggacoggtactggtctccaggagcgagggggtgggtgccag	1970
Db	1861	GTTACTGCACAGGTTCCGGAGCCGGTACTGTCTCCAGGACGAGGGGGCCAGGGGTGCCAG	1920
Qy	1971	ggaggtggctccactccagctctcctctccggtccggtttctgtctcaccctacatc	2030
Db	1921	GGAGGTGGCTCTTACCCAGCTTCCTCTCTCCCTTCCTGGTTCCTGTCTCCACACCTACATC	1980
Qy	2031	cccaaccctctttgcccagcagggcatgacccctccctcgccctgacagccctcc	2090
Db	1981	CCCGCCACCTCTTTGGCCACGAGGCGACGCCCTTCCCTTGCCCTGACAGCCCTCC	2040
Qy	2091	acctgctatgctactaggtcccccagtgccctccactggactctcggtcgccgctgactg	2150
Db	2041	ACCTGCCTATGCTACCTAGGCCCTCTGCTCTCCACTTGACTCGGCTGTGCTGGCTGCTC	2100
Qy	2151	ttctgctctgcccctggc---ggcgctctgagagagagagagatctctctcaccagccc	2207
Db	2101	TTCTGCTGTGTCTCCGGCCATGGCACTGTGAGAGAGAGAGTGTCTCTCACCCAGGCC	2160
Qy	2208	tgagctccctgttaactgatggatatctcacttcagcacatcttcccacgcccgcgc	2267
Db	2161	AGTGCTCC-----TGTCATTTCAGACATCTTCCACAGCTGCCAGC	2202
Qy	2268	cgagccctctggtgtgctgctgaggaacagggctgagtaaggggggccaggaagccg	2327
Db	2203	TGCTCTTTGGTGTG-----TCCCGGAGGAACAGGCTAAGCAAGGGGCCAGGAAGTTG	2254
Qy	2328	tcaagagaacaggggttaacgatctgtctagagctgcccctgcccctcttctgaaactg	2387
Db	2255	TCCAGAGGACAGGGGCTTAATGAGCTGTCTGAGAGTGCCCTGCCCTGCTCTGAACACTG	2314
Qy	2388	ctctc---cacaagcacagacaagctccctctccctcgatcagccagccagacttga	2445
Db	2315	CTTTCCACACAGCACGGGCAAGTCCCTTTACCTTTGGATCAGCAAGCCAGACTTGGGA	2374
Qy	2446	gctctgacaaggaactttccttggtgagagaccaggaagaa-----c	2484
Db	2375	GCTTTGACAGGACGTTCCCGGGAACGACCAACGACCAACATATAACAAGGC	2434
Qy	2485	aaatagaaacgcacatcactctaccccgtgaccagaggtgctgctcactgccc	2544
Db	2435	ACAGAGAAGTGGCCACAGCCTTCCACCCACGACACAGAGACTGGCCTGGCCTCACTGCT	2494
Qy	2545	ttaaggaaactgatgtctctctctcttgaacttgggtggggaacccccaccccaaaag	2604
Db	2495	TTCAAGGACACAGATGTCTACCTCTCTTGAACCTTTGGTGGGGAAACCCACCCAAAG	2554
Qy	2605	ccccctgttagtctcttggcaatccccctccctcactcctcaggtgggactagagt	2664
Db	2555	CCCCCTTTGTAGTCTC---TTGGCAATTTCTTCCCTCTCACTCTCTCAGGTGGGGCTAGAGT	2613
Qy	2665	aagacagaccaggtgagcgtgggtgtctctctcattctcaagactctctcattgtatc	2724
Db	2614	AAGTCTGACATCCCTCCCATTTCTCAAGACTCTC-----TCTCTTTTCATTGGGTAC	2664
Qy	2725	acaataccacagtgacctctgtggtgtgtctctctctctctctctctctctctctctct	2784
Db	2665	CCTGTACCCAGTGCCTCTGCGTGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2706

FNDKNNLWMSMPGVNNGLSLTLRTEQNDIFLLSTVTGARVWVHODEPAMDDG
GNLRPGVTSISMRKEALDSLGGNYDGTGSDVPVKNLYPSKYTOVCVTHSCFOE
NMILKGGCAVIFVPPKPGYDFCDYRKQSWGVCYKLOGAFSLDSLGCFSKRGKPCSV
INYKLSAGYSRPSVKSQDIWEMLSLONNYINNRKNGVAKLNFYKELNXTKNS
PSVTWVLSLJNSQMSLWFGSVLSVWEMAEILFDLLVTLMLLRRFRSRYWSPGR
GARGAREVATFASSPFSPFCPHPTSPPLPQQGMTPTPLALTAPPPAYATLGPSAPP
LDSAAPDCSACALAAI"

BASE COUNT 480 a 699 c 545 g 476 t
ORIGIN

Query Match 70.5%; Score 2196.8; DB 12; Length 2200;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	47	gtcccgagacgcccattctgccttaacgctaatgatgtgagccacacacagagcccctg	106
DB	1	gtcccgagacgcccattctgccttaacgctaatgatgtgagccacacacagagcccctg	60
QY	107	agctcaacattgacctagaccttcaagcctccaaactgcctcaaggggtccatgaaggca	166
DB	61	agctcaacattgacctagaccttcaagcctccaaactgcctcaaggggtccatgaaggca	120
QY	167	accaattcaaggagcaagacccttgtctctcagcccatgcaagagactgggaagggg	226
DB	121	accaattcaaggagcaagacccttgtctctcagcccatgcaagagactgggaagggg	180
QY	227	acaaactgaagagcagggcctggcccggaacccctcagcaaccccgagccacccaggg	286
DB	181	acaaactgaagagcagggcctggcccggaacccctcagcaaccccgagccacccaggg	240
QY	287	aggagagcactgattgaattccaccgctcctaccggagctctccagttctcttgca	346
DB	241	aggagagcactgattgaattccaccgctcctaccggagctctccagttctcttgca	300
QY	347	acaaacacacattccacggggccatccgctggtgtgctccaaacacaaacgcgatga	405
DB	301	acaaacacacattccacggggccatccgctggtgtgctccaaacacaaacgcgatga	360
QY	407	cggcctctggcggtgctgtgctgtgaccttggcagctgactggcagctgcct	465
DB	361	cggcctctggcggtgctgtgctgtgaccttggcagctgactggcagctgcct	420
QY	467	tgtgttcgagagtagctcagctaccagctgagcctcaacatcaacctcaattcagaca	526
DB	421	tgtgttcgagagtagctcagctaccagctgagcctcaacatcaacctcaattcagaca	480
QY	527	agctgtcttcctgcctgactgtgtgaccccttaactcctacagatacactgaaatta	586
DB	481	agctgtcttcctgcctgactgtgtgaccccttaactcctacagatacactgaaatta	540
QY	587	aagagagctggaagagctggaacgcatcacgagagcagccttttggactgtacaaat	646
DB	541	aagagagctggaagagctggaacgcatcacgagagcagccttttggactgtacaaat	600
QY	647	acaactcttcacactcgagcctggggcgagccgagcctcccgagcctcctg	706
DB	601	acaactcttcacactcgagcctggggcgagccgagcctcccgagcctcctg	660
QY	707	gtgtcttcggcagccctcagcctgcgactccacccctcagcctcactccggcgca	766
DB	661	gtgtcttcggcagccctcagcctgcgactccacccctcagcctcactccggcgca	720
QY	767	cggcgagcaggggtcttcagcgtacgcgaacaaatcccccaagtggacgggaagact	826
DB	721	cggcgagcaggggtcttcagcgtacgcgaacaaatcccccaagtggacgggaagact	780
QY	827	ggaagatcgacttccaaactgtgaacgaacaaatcagactgtttctacagacatact	886
DB	781	ggaagatcgacttccaaactgtgaacgaacaaatcagactgtttctacagacatact	840
QY	887	cctctgggtgtagtgagtgagtggtgtaccgcttccattacatacaactctgtcca	946

DB	841	CTCTGGGGTGGATGTCAGTGGGGAGTGGTACCCTTCCATTACATCAACATTCTGTCCA	900
QY	947	gactgtcggaacacctgcccgcctctagaggaagaagccctgggcaacttcatctacact	1006
DB	901	GACTGTGCGGACACTGCCCGCTCTAGAGGAAGAAGCCCTGGGCAACTTTCATCTCACCT	960
QY	1007	gtcgcttcaaccaggccccctgcaaccaggcgaattattccaagtctccaccacccccatgt	1066
DB	961	GTGCTTCAACAGGCCCTCCGCAACAGGCGGAATTATTCGAAGTTCACACCCCATGT	1020
QY	1067	acgggaactgtacactttcaatgacagaacaaactccaatctctggatgtctctccatgc	1126
DB	1021	ACGGGAAGTGTACACTTTCAATGACAAGAACAACCTCCAATCTCTGGATGTCTCTCCATGC	1080
QY	1127	ctggagtcacaatggtttgtccctgacactgcacagagcagaatgacttcatcccc	1186
DB	1081	CTGGAGTCAACAATGTTGTCCCTGACACTGCGCAGAGCAGAAATGACTTTCATCCCTC	1140
QY	1187	tgtgtccacagtgacgggggcccaggggtggtgctgctggtgcagatgagcctgccttta	1246
DB	1141	TGCTGTCCACAGTGACGGGGCCAGGGTGATGGTGCATGGTCAGGATGAGCCTGCCCTTA	1200
QY	1247	tgatgtgtgtggttcaacttggagcctggcgtgaggaacctccatcagtatgagaagg	1306
DB	1201	TGGATGATGGTGGCTTCAACTTGAGGCTGCGGTGGAGACCTCCATCAGTATGAGAAGG	1260
QY	1307	aagccctggacagcctcgagggaattacggcgactgtactgagaatggtagcatgtcc	1366
DB	1261	AAGCCCTGGAGAGCCTCGAGGAAATTTACGGCGACTGTACTGAGAATGGTAGCATGTCC	1320
QY	1367	cgttcaagaaccttacccttccaaagtatacacagcaggtgtgcatcactcctgcttcc	1426
DB	1321	CGGTCAAGAACCTTTACCTTCCAAAGTATACACAGCAGGTGTGCATTCACCTCTCTCTCC	1380
QY	1427	aggagaacatgatacaagaagtgtggtgtgctacatcttctacccctaaagccccaggag	1486
DB	1381	AGGAGAACATGATCAAGAAGTGTGGCTGTGCTACATCTTCTACCTTAAGCCCAAGGAG	1440
QY	1487	ttgagttctgtgactaccgaaagcagagcctcctgggctattgctattataaaactgagg	1546
DB	1441	TTGAGTTCTGTGACTACCGAAAGCAGAGCTCCTGGGCTATTGTCTATTATAAATGTCAGG	1500
QY	1547	gcgcttctcctgtgacagcctgggctgttctccaaagtgtcgaagccttctgtgagtg	1606
DB	1501	GGCCTTCTCCTTGGACAGCCTGGGCTGTTTCTCCAAAGTGTCCGAAGCCTTGTAGTGA	1560
QY	1607	tcaactaaactctctgcgggtactcaoggtggccatctgtgaagtcaccagattgga	1666
DB	1561	TCAACTACAAACTCTCTGCGGCTACTCAGGGTGGCATCTGTGAAGTCCCGAGATTGGA	1620
QY	1667	tcttcgagatgtgtccttgcagaaacattacactattacaaacaaagaaagaggttg	1726
DB	1621	TCATTGAGATGCTCTCTTGCAGAACAAATTACACTATTAAACAAAGAAAGAGAGTTG	1680
QY	1727	caagctcaacatcttctcaaggagctgaactataaaactaatcggagctcctctctg	1786
DB	1681	CAAAGCTCAACATCTTCTTCAAGGAGCTGAACATAAATAAATTCGAGTCTCTCTCTG	1740
QY	1787	tcacgatgtgtcagcctcctgtccaaacctgggagccagtgagcctgtgtgttggctcgt	1846
DB	1741	TCAGATGGTCAGCCTCTCTGTCCAACCTGGGCAGCCAGTGGAGCTGTGTGTGGTTCGT	1800
QY	1847	ccgtgtctctgtgtgtgagatggcgaagctcatctctgaacctccttggtatcacacttc	1906
DB	1801	CCGTGTCTCTGTGTGGTGGAGATGGCGAGCTCATCTTCGACCTCTGTCTCATCACATTC	1860
QY	1907	tcatgtcttacgcctgttccggagcagtgactgtctccagagcaggggcccagggtg	1966
DB	1861	TCAATGTCTACGCCGGTTCGGAGCCGGTACTGTCTCCAGGACGAGGGGCCAGGGGTG	1920
QY	1967	ccagggaggtggcctccactccagcttctccttcccgctcccggttctctgtcctcaccta	2026
DB	1921	CCAGGAGGTGGCCTCCACTCCAGCTTCTCTTCCCGTCCCGTTTCTGTCTCTCACCTTA	1980

Db	961	GTGCTTTAAACAGGCCCCCTGCAACCGAGGGGAATTAATTCACAGTTCCACCAACCCCAATGT	1020
Qy	1067	acgggaactgctacacttttcaatgacaagaacaactccaatctctgtagtgtccctccatgc	1126
Db	1021	ACGGGAACTCGTACACTTTCAATGACAAGAACAACTCCAATCTCTGGATGTCTCTCCATGC	1080
Qy	1127	ctggagtcacaatggttgttccctgacactgccccagagcgacagatgaacttcaatcccc	1186
Db	1081	CTGGAGTCAACAATGGTTGTCCCTGACACTGCGCACAGACAGAAATGACTTTCATCCCCC	1140
Qy	1187	tgctgtccacagtgacgggggcagggtgaagtgcattggtcaagatgagccctgccttta	1246
Db	1141	TGCTGTCCACAGTACGGGGGGCAGGGTGAATGGTCATGCGTCAAGATAGCCTGCCCTTA	1200
Qy	1247	tgatgatggtgctcaacttgagcctggcgtggagacctccatcagtatgagaaagg	1306
Db	1201	TGATGATGCTGGCTTCACTTGAGGCCCTGGCGTGGAGACCTCCATCAGTATGAGAAAGG	1260
Qy	1307	aagccctggacagcctcggaggaaattacggcgactgtactgagaatggtagcagatgcc	1366
Db	1261	AAGCCCTGGAGACCCTCGGAGGAAATTAACGGCGACTGTACTGAGAATGGTAGCGATGCC	1320
Qy	1367	cggctcaagaacctttacccctccaagtatacacagcagggtggaattcaactcctgcttcc	1426
Db	1321	CGGTCAAGAACCTTTACCTTCCAAAGTATACACACAGCGTGTGCATTACACTTCCTGCTTC	1380
Qy	1427	aggagaacatgatcaagaagtggcgtgacctacatcttaccctaaagccaaagagag	1486
Db	1381	AGGAAACATGATCAAGAGTGTGGCTGTGCCCTACATCTTTACCTTAAGCCCAAGGAGG	1440
Qy	1487	ttgagttctgtgactaccgaagcagagctcctggggctattgctattataaactgcagg	1546
Db	1441	TTGAGTTCTGTGACTACCGAAAGCAGAGCTCTGGGGCTATTGCTATTATAAATGCAGG	1500
Qy	1547	gcgccttctccttggaagcctgggctgtttctccaagtgcggaagcctgtagtgtga	1606
Db	1501	GCGCCTTCTCCTTTGAGACGCTGGCGTGTTTCTCCAAGTCTCGAAGCCTTGTAGTGTGA	1560
Qy	1607	tcaactacaaactctctgcggctactcacgtggccatctgaaagtcgccagatigga	1666
Db	1561	TCAACTACAAACTCTCTGCCGGCTACTCAGGTGGCCATCTGTGAAGTCCCAGGATTTGGA	1620
Qy	1667	tcttcgagatgctgctcttgagaacaattacactattaaacaaagaagaacaggagtgtg	1726
Db	1621	TC TTCGAGATGCTGCTCTTGAGAACATTAACACTATTAAACAACAAAGAAACGAGCTTG	1680
Qy	1727	caagctcaacatctcttccaaggagctgaactataaactaattcggagtcctcctctg	1786
Db	1681	CAAAAGCTCAACATCTCTTCAAGAGCTGAACATATAAACTAAATTCGGAGGTCTCCTTCG	1740
Qy	1787	tcagatggtcagccctcctgtccaactgggcagccagtcagcctgtggtttgagctcgt	1846
Db	1741	TCAGATGGTCAGCCCTCTGTCCAACCTGGGAGCCAGTGGAGCCTGTGGTTGGCTCGT	1800
Qy	1847	ccgtgctctctgtgttgagatggcggaagtcattctcgacctcctgtgtcatcacatctc	1906
Db	1801	CCGTGCTCTCTGTGTGGAGATGGCGGAGCTCATCTTCGACCTCCTGTGTATCATACTTC	1860
Qy	1907	tcatgtctacgcgggttcaggagccggtactggtgtccaggaagggggccagggggtg	1966
Db	1861	TCATGCTGTCTACGCCGTTCCGGAGCGGTACTGTGCTCCAGGACGAGGGGGCGAGGGTG	1920
Qy	1967	ccagggaagtggcctccaactcagcttctcctcccggtccggtttctgctcaacctta	2026
Db	1921	CCAGGGAGGTGGGCTCCACTTCACGTTCCTCTCTCCCGTCCGGTTTCTGTCTCTCACCCCTA	1980
Qy	2027	catccccaccaactctttgccccacagcgggcagtcagccccctccctggccctgacagccc	2086
Db	1981	CATCCCCACCACTCTTGTGCCCCAGCAGGGATGACCCCTCCCTGGCCTGTGACAGCCC	2040
Qy	2087	ctccactgcctatgtactctaggccccagtcgccccctccactggactctgcggtgcgctg	2146
Db	2041	CTCACCTGTCTATGTACTCTAGGCCCCAGTGGCCCTCCACTGTGACTCTGGCGGCTG	2100

Qy	2147	actgttctgcctgtgccttcgctgcctctgagagagagagagatcctctcaaccagcc	2206
Db	2101	ACTGTTCTGCCTGTGCCCTGGCGGCGTCTGAGAGAGAGAAGGATCCTCTCACCCAGCC	2160
Qy	2207	ctgaactcccttgtaaacatgatgatctcacttcagca	2246
Db	2161	CTGAGCTCCCTTGTAACATGATGGATATCTCACTTTCAGCA	2200
RESULT	6		
HSLASNA			
LOCUS	HSLASNA	3151 bp	mRNA
DEFINITION	H.sapiens mRNA for lung amiloride sensitive Na+ channel protein.		01-FEB-1994
ACCESSION	X76180		
VERSION	X76180.1	GI:452649	
KEYWORDS	Na+ channel; Na+ channel protein.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
AUTHORS	Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 3151)		
JOURNAL	Barbry, P.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (19-NOV-1993) P. Barbry, CNRS, IPMC 660 Route des		
TITLE	Lucioles, 06560 Sophia Antipolis, FRANCE		
JOURNAL	2 (bases 1 to 3151)		
REFERENCE	Voilley, N., Lingueglia, E., Champigny, G., Mattei, M. G., Waldmann, R.,		
AUTHORS	Lazdunski, M. and Barbry, P.		
TITLE	The lung amiloride-sensitive Na+ channel: biophysical properties,		
JOURNAL	pharmacology, ontogenesis, and molecular cloning		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 91 (1), 247-251 (1994)		
FEATURES	94105144		
source	Location/Qualifiers		
	1..3151		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="12"		
	/tissue_type="lung"		
	/cell_type="epithelial"		
	/dev_stage="adult"		
	100..2109		
	/codon_start=1		
	/product="Na+ channel protein"		
	/protein_id="CAAS3773.1"		
	/db_xref="GI:452650"		
	/db_xref="SWISS-PROT:P37088"		
	/translation="MEGKLEEQDSPSPQSTPGMLKMKREEQGLGPEAPAPQOPTAE"		
	EEALTEFHSRYELFEPSKYQOVCIHSCFOESMIKEGCAVIFYPQNVCEYCDYRKH		
	GLLGEVEFSPVSLINLNSDKLVFPATVICTLNPYFYPETKELELDRTTEQLD		
	LYKYSSTTTLVAGSRERDLGTLPHLQLRLVPPPHGARRARSVASSLRDNNPQVD		
	WKWDKGFQJCNKSDCFYQYSSGVDVAFREYRFHYINILSRPLTEFLSEEDTUG		
	NIETAFACNVSCNQAYSHFHPHMGNCYTFDNKNSLNMMSMPGNNGLSLMLRA		
	EQNDIFILLSVTGARVMVQDEPAMDGDGFLNLRPGVETSIEMRKETIDRLGDDGYD		
	CTKNGSDPVENLYPSKYQOVCIHSCFOESMIKEGCAVIFYPQNVCEYCDYRKH		
	SSWGCYCYKLDVDSLDHLCGFTCKRXCPSVTSQLSAGYSRMPSVTSQEWVQMLSR		
	QNNTYVANKRGVAKNVFKELNYKTNSESPSVMTVLLSNLGSQSNLFGSSVLSY		
	VEMAELFVDDLVIMFLMLLRFRSRYWSPGRGGGAQEVASTLASSPPSHFCFHPHMS		
	SLSGPFPSPALTPATVPAYATVLPREPSPGGSGASSSTCPLGCP"		
BASE COUNT	995 c 803 g 676 t		
ORIGIN			

	Query Match	49.2%	Score 1532.2	DB 9	Length 3151
	Best Local Similarity	73.3%	Prod. No. 0		
	Matches 2254	Conservative	0	Mismatches 678	Indels 145
				Gaps	17
yy	151	gggtccatgaagggaacaaattcaaggagcaagacccttgctctccagccga	210		
bb	94	GGTCTCATGGAGGGGAACAAAGCTGGAGGAGCAGGACTTAGCCCTCCACAGTCCACTCCA	153		
yy	211	ggactgggaagggggacaaacgtgaagagcagggcctgggcccggaaacctcagcaccc	270		

Db 154 GGGCTCATGAAGGGGAACAAGCGTGAGGAGCAGGGGCTGGGCCCCCAAGACTCGCGCGCCC 213
Qy 271 cggcagccaccgagagagagagagcactgattgaattccaccgcctcctaccggagagctc 330
Db 214 CAGCAGCCACCGGAGGAGGAGGCCCTGATCGAGTTCCACCGCTCTCCAGAGAGCTC 273
Qy 331 ttccagtttcttctgcaacaacaccaccattccacggyggccatccgcctggtgctgccaaa 390
Db 274 TTTCGAGTTCTTCTGCAACAACACCACCATCCACGGCGCCATCCGCTGTGTGCTCCAG 333
Qy 391 cacaaccgcataagaacggcctcttggcggtgctggtggtggtgacctctggcagtatg 450
Db 334 CACAACCCGATGAAGACGGCTTCTGGCAGTGTCTGGCTCTGACCTTTGGCATGATG 393
Qy 451 tactggcagttcgcttgcgttcgaaggagtagacctcagctaccctacccagtgacctcaacatc 510
Db 394 TACTGGCAATTTCGGCTGCTTTTCGGAGAGTACTTTCAGCTACCCCGTCAAGCTTCAACATC 453
Qy 511 aacctcaattcagacaagctgtcttccctgcctgcctactgtgtgaaccttaactcttac 570
Db 454 AACCTCAACTCGGACAAGCTCTGTTCCCGCAGTGACCATCTGCACCCCTCAATCCCTAC 513
Qy 571 agatcacactgaattaaagagagctggaagagctggacgcatcaccagagacagcgtt 630
Db 514 AGGTACCCCGGAATTAAGAGAGAGCTGGAGAGCTGGACCGCATCACAGAGCAGAGCTC 573
Qy 631 ttgacttgtacaatatacaactcttctacactcgcaggtcgccaggtggggccgcagccgagc 690
Db 574 TTGTGACCTGTACAAATACAGCTCTTCACCACTCTC- ----GTGGCCGGCTCCCGCAGC 627
Qy 691 tcccgcgacctcctgggtgtcttccgcgaacccctgcagcgctgcgcactccacctcg 750
Db 628 CGTCGCGACCTCGGGGGACTCTGCCGCACCCCTTTCAGCGCTGAGGGTCCCGCCCCG 687
Qy 751 cctactccggcgcaacggcgagcgagcggtgttccagcgtacgcgacaacaatccccaa 810
Db 688 CTCACGGGGCCCGTCGACGCCGTAGCGTGGCTTCAGCTTCGGGACAAACAACCCCCAG 747
Qy 811 gtggcaggaagactggaagatcggtctccaactgtcaaccagacaacaatcagactgt 870
Db 748 GTGGACTGGAAGACTGGGAAGATCGCTTCCAGCTGTGCAACCAAGAAATCGGACTGC 807
Qy 871 ttctaccagacatactcctctgggtggatgagtcagtgagggagtggaccgcttccattac 930
Db 808 TTCTACACACATCTCATCAGGGTGGATGCGGTGAGGGAGTGGTACCGCTTCCACTAC 867
Qy 931 atcaacattctgtccagactgtcgacacactgcgcctctctagaggaagaagccctgggc 990
Db 868 ATCAACATCTGTTCGAGGCTGCGCAGAGACTCTGCCATCCCTGGAGGAGGACACGCTGGGC 927
Qy 991 aacttcacttccactgtcgttccaccagggccctcgcaaccaggcgaattatccaag 1050
Db 928 AACTTCATCTTCGCTTGGCCGTTCACACAGGTCTCTCGCAACGAGCGAATTACTCTCAC 987
Qy 1051 ttccaccaccctatgacgggaactgctacactttcaatgacaagaacaactccaatctc 1110
Db 988 TTCCACCACCCGATGTATGGAACATGCTATATTTCATGACAAAGAACTCCAACCTC 1047
Qy 1111 tggatgtcctccatgctcgtgagtaacaatggttgtcctctgacactgcgcagagcag 1170
Db 1048 TGGATGTCTTCCATGCTGGAAATCAACAACGGTCTGTCCCTGATGCTGCGCGCAGAGCAG 1107
Qy 1171 aatgacttccctctgcttccacagtgacggggccaggggtgatggtgatgtcag 1230
Db 1108 AATGACTTCAATTCCTTCCTGCTTCACAGTGACTGGGGCCGGGTAAATGTGACGGGCAG 1167
Qy 1231 gatgagctgcctttatgatgatggttggtctcaacttgagccttgagcctggagaccttc 1290
Db 1168 GATGAACCTGCCCTTATGGATGATGGTGGCTTTAACTTCGGCCCTGGCGTGGAGACTCC 1227
Qy 1291 atcagtatgagaagaagccctggacagcctcggaggaaattacggcgactgtactgag 1350
Db 1228 ATCAGCATGAGGAAGAAACCTGGACAGACTTGGGGGGGATTATGGCGCAGTGCAACAAG 1287

Qy 1351 aatggtagcgtatcccggtcaagaacctttacccttcccaagtatacacagcagggtgtgc 1410
Db 1288 AATGGCAGTGATGTTCTGTTGAGAACCITTAACCTTCAAGTACACAGCAGGTTGTGT 1347
Qy 1411 attcactcctgtctccagagagaacatgatacaagaagtgtggtctgtoctacattctctac 1470
Db 1348 ATTCACCTCCTGCTTCCAGGAGAGCATGATCAAGGAGTGTGGCTGTGCTACATCTTCTAT 1407
Qy 1471 cctaagcccaagggagtgtgagttctgtactaccgaaagcagagactcctgggctattgc 1530
Db 1408 CCGCGCCCCAGAACCTGGAGTACTGTGACTCAGAAAAGCACAGTTCTCTGGGGTACTGC 1467
Qy 1531 tattataaactgcagggcgcttctctctgtgacagcgtggctgtttctccaagtgtcgg 1590
Db 1468 TACTATAAGCTCCAGGTTGACTTCTCTCAGACCACTGGGCTGTTTCACCAAGTGCCTG 1527
Qy 1591 aagcctttagtgtgatcaactacaactctctgcgggtactcaagggtggcctatctgtg 1650
Db 1528 AAGCCATGCAGCTGACCACTACAGCTCTCTGCTGTTACTCAGATGGCCCTCGGTG 1587
Qy 1651 aagtcaccagattggatcttcgagatgctgtccttgcagaaacaaattacactattaacaac 1710
Db 1588 ACATCCCAGGAATGGGTCTTCCAGATGCTATCGGACAGAACAAATTACACCGTCAACAC 1647
Qy 1711 aaaaagaaacggagtgtcaaaagctcaacatcttctcaaggagctgaactataaaactaat 1770
Db 1648 AAGAGAAATGGAGTGGCCAAAGTCAACAICTTCTTCNAGGAGCTGAACACTACAAAACCAAT 1707
Qy 1771 tggagtctcctctctgtcacgatgtgtagcctcctctccaacctgggagccagtgagc 1830
Db 1708 TCTGAGTCTCCCTCTGTACGATGGTACACCTCTGTGCCAACCTGGSCAGCCAGTGGAGC 1767
Qy 1831 ctgtggttggctgcgtcctgtctctctgtgtgtagagatggcgagcgtcatcttcgaacct 1890
Db 1768 CTGTGTTTGGCTCCCTCGGTGTTGTGTGTGGTGGAGATGSGCTGAGCTGCTTTGACCTG 1827
Qy 1891 ctggtcatcacactctcatgctgcacgggttccgagcggtactggtgtactggtctccagga 1950
Db 1828 CTGCTCATCATGTTCTTCATGCTGCTCCGAGGTTCCGAAGCCGATACTGGTCTCCAGGC 1887
Qy 1951 cgagggggcagggtgccagggaggtggctccactccagcttccctccttcccgctccctg 2010
Db 1888 CGAGGGGCGAGGGTGTCTCAGGAGGTAGCTCCACCTGGCATCCTTCCCTCTCTCCAC 1947
Qy 2011 ttctgtcctcaacctacatccccacacacttcttggccccagcagggatgacacctccc 2070
Db 1948 TTCTGCCCCACCCCATGT-----CTCTGTCTTGTGCCAGCCAGGCCCTGTCCCTCT 2001
Qy 2071 ctggccctgacagccctccacctgcctatgctactcttagccccagtgccccctcaactg 2130
Db 2002 CAGCCTTGACACGCCCTCCCTCTGCTATGCCCACCTGGGGCCCCGCCCATCTCCAGGG 2061
Qy 2131 gactctggcgctgactgttctgctgtgccttggcgctgagagagagagagag--aa 2188
Db 2062 GGTCTCTGAGGGGCCAGTTCTCTCCACTGTCTCTGGGGGGCCCTGAGAGGGAAGAGA 2121
Qy 2189 ggaatcctcaaccagccctgagctcc-----ctgttaaaactgatgatactcaactcag 2244
Db 2122 GGTTCCTCACCAACGAGCAGATGCTCTGTGTGGAGGGTGTGGCCCTTGGCAAGATTG 2181
Qy 2245 cacatcttccccagggcccccgccgagc-----ctctgtgtgtgcttccctcagggaa 2299
Db 2182 AAGGATGTGAGGGCTTCTCTCTCAGAGCCCCCAACTGCCCTTGATGTGTGGAGGGAA 2241
Qy 2300 caggctgagtaaggggcccaagcgttccagagagacaggggctaacg-atctgtcctcag 2358
Db 2242 GCAAGATGGGTAAAGGCTCAGGAAGTGTCTCCAAAGAACAGTAGCTGATGAAGCTGCCAG 2301
Qy 2359 agctgccttgcctctgcttcttgaaac-----tgcctt 2391
Db 2302 AAGTGCCTTGGCTCCAGCCCTGTACCCCTTGTACTGCTTGTGAACTCTGTGGTTTCCCC 2361

QY 2392 ccacaagcacagacaagtccctctctctccctcgatccagccagccagccagcttgagctctg 2451
Db 2362 ACCCAACTGCGGCTAAGTCTCTTTTCCCTTGGATCAGCAAGCAAACTTGGAGCTTG 2421
QY 2452 acaagaaacttctctggagagacaaagaaacaaatagacagccatcagctctctacc 2511
Db 2422 ACAGGAACCTTCTTAAGAAACCGCTGATACCAAGCAACAAACACCAAGGCTACAG 2481
QY 2512 ccgctgaccagag-----ctgcccagag-----ctgcccagag 2536
Db 2482 CAGCATGACAGGGTTCTTCCCGCAGGAGCGGTTAAGCCAGCCCGCAGCTGGCTGGCC 2541
QY 2537 tcaactgc-cttcaggagacactgatctctctctctctctctctctctctctctctctctct 2595
Db 2542 ACAGTCTCTCCAGTACGACAGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2601
QY 2596 acccaaaagccct 2655
Db 2602 ACCCAAGGCCCT--TTGTTACTTAAAGCAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2659
QY 2656 gactagagtaagacagaccaggttagcctg--gtgtctctctctctctctctctctctctct 2713
Db 2660 GGCTAGAGCAGACCGGGTAAAGAGGACACCGAGGCTCTCTCTCTCTCTCTCTCTCTCTCT 2719
QY 2714 tcaattgatcacatacccccagct 2764
Db 2720 GCCTCAGAGGACGCGCCGCGACCTCTGCGCTGTCTTTCATACCTCTACATGCTCT 2779
QY 2765 -----tgttctcagcctggaagctctctctctctctctctctctctctctctctctctct 2816
Db 2780 GCTTGAGATATTTCTCAGCCTGAAAGTTTCCCAACACCATCTGCCAGAGAACTC--CTA 2836
QY 2817 tgcactcttagaacaactgcccagaagcaccattctctctctctctctctctctctctctctct 2876
Db 2837 TGCATCCCTTAGAACCTGCTC--AGACACCATTTCTTTGTGAACCTCTCTGCCACATC 2894
QY 2877 atgtctctcccaagaatg-----tcgcccctctctctctctctctctctctctctctctct 2931
Db 2895 TTGCTTCCCAAAATTTGATCACTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2954
QY 2932 tatg-----ttcacattct 2965
Db 2955 TCTGCTGGAGTGTGCTGTGACCATACTTCTTGTGACATTTGCTGCTCTCTCTCTCTCTCT 3014
QY 2966 taaactgagctctctgtgagtcagga-tgagatttctctctctctctctctctctctctctct 3024
Db 3015 TAGACTTAAGTGCTTGGCTGAGGACTGAATCTTGGCCGTTTATGTATGCTTCCATGT 3074
QY 3025 ctgcccagctctccacttggagcggtaggcaggtactcaataaagcttctctctctctctct 3084
Db 3075 CTAGCCCATCATCTGCTTGGAGCAAGTAGCAGGAGCTCAATAAATGTTTGTTCATGA 3134
QY 3085 aaaaaaaaaaaaaa 3101
Db 3135 AAAAAAAAAAAAAA 3151

RESULT 7
HUMSODIUM 2334 bp DNA PRI 29-SEP-1994
LOCUS Human kidney aniloride-sensitive sodium channel, complete cds.
DEFINITION L29007
ACCESSION L29007
VERSION L29007.1 GI:493125
KEYWORDS sodium channel
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2334)
REFERENCE McDonald,F.J., Snyder,P.M., McCray,P.B. Jr. and Welsh,M.J.
AUTHORS Cloning, expression, and tissue distribution of a human
TITLE aniloride-sensitive Na+ channel
JOURNAL Am. J. Physiol. 266 (Pt 1), L728-L734 (1994)

MEDLINE 94295729 Location/Qualifiers
FEATURES
source 1..2334
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="Kidney"
83..2092
/partial
/codon_start=1
/product="Na+ channel"
/protein_id="AAA21813.1"
/db_xref="GI:493605"
/translation="MEGKLEQDSSPPSTPGIMKGNKREQGLGPEPAAPQPTAE
EEALIEFHSYRELFFFCNNTIHAIRLVCVSOHNRMTAFWVWLWLCFQGMWQF
GLLFGEFHSYPSVLINILNSDKLVFPAVITCLINPVRYPEIKELDELDRITQTLFD
LYKSFTILVAGSRSRDLRGLTLPQLRLVPPPHGARRARSVASSLRDNPQVD
WDMKLGQLCNQNSCDFYQTSYSSGVDAVREWRHYINILSRPELTLSLEEDTIG
NKDWKIFQVSCNANYSHFHPMYGNCYFNDKNNLSMMSMFGINGLSLMURA
EQNDFIPLISTVTGARVMHGDPAFMDGDFNLPGVETSIEMRKETLDVLDGDDYD
DCTKNGSDPVENLYPSKYTOQVCIHSCFOESMIKCGCAYIFYPONVETCDYRKH
SSNGCYCYKLOVDFSSDHLGCTKCRKPCSVTSYQLSAGYSRPSVTSQSMVQMLSR
ONNYTNNKRGVAKYNIFFKELNYTNSPSPYMTWTLNLSNLGSQLSFWGSSVLSY
VEMAEIVFDLLVIMFLMLLRFRSRYNSPGRGGAGAEVASTLASPPSHFCHPMSL
SUSQGPAPSPALTPAPPAYATLGPSPGSGASSTCPLGP"
BASE COUNT 486 a 739 c 632 g 477 t
ORIGIN

Query Match 46.1%; Score 1436.8; DB 9; Length 2334;
Best Local Similarity 82.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 357; Indels 12; Gaps 2;

QY 151 gggtccatgaaggcaaccattcaaggagcaagaccctctctctcaccgccatgcaa 210
Db 77 GGCTCTCATGGAGGAAACAAGCTGGAGGACAGACTTAGCCCTCCACAGTCCACATCCA 136
QY 211 ggactgggggaagggaacacgtgaagagcagcgctggccggccgaacctcagcacc 270
Db 137 GGGCTCATGAAGGGGAACAACGGTGAAGAGCAGGGGCTGGGCCCCCGAACCCTGCGCGGCC 196
QY 271 cggcagccccagggaggaggagggcactgattgaattccaccgctcctcaccggagctc 330
Db 197 CAGCAGCCACGGCGGAGGAGGCGCTGTATGAGTTCCACCGCTCTCTACCGAGAGCTC 256
QY 331 ttccagttctctgcaacaac 390
Db 257 TTCAGTCTCTTCGCAACAAC 316
QY 391 cacaaccgcatgaagacggccttctggcggtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 450
Db 317 CACAACCGCATGAAGACGGCTTCTGGSCAGTGTGTGGCTCTGCACCTTTGGCATGATG 376
QY 451 tactggcagttcgcctgtctgttcgagaggtacctcagctaccacagtgagcctcaaatc 510
Db 377 TACTGGCAATTCGGCTGCTTTTCGAGAGTACTTACGCTACCCCGCTCAGCTCAACATC 436
QY 511 aacctcaattcagacaagctgttctcctcctcctcctcctcctcctcctcctcctcctccttac 570
Db 437 AACCTAACCTGGGACAAGCTGCTTCTCCCGCAGTGAACCATCTGCACCTCAATCCCTTAC 496
QY 571 agatcacactgaaataaagagagagctggaagagctggaccgcatcacggagcagacgctt 630
Db 497 AGGTACCCGGAAATTAAGAGGAGCTGGAGAGCTGGACCGCATCACAGACGACGCTC 556
QY 631 ttgactgttacaataacaaactcttctacactcgcaggctggcgccgagcgcgcgcgcgcgc 690
Db 557 TTTGACCTGTACAAATACAGCTCTTCCACTCTCTCGTGGCGGCTC-----CCGCAGC 610
QY 691 tcccggaactcctcgtgtgttctccgcacccccctgagcgcctgcgcgcctcgcacacctccg 750
Db 611 CGTCGACACCTGCGGGGGACTCTGCCGACCCCTTGACGCGCTGAGGGTCCCGCCCGCC 670
QY 751 ccctactccgc 810

GFTKCRKPCSVIIYKLSASYQWPSATSQDWMFQMLSRNNYTIKNKRGDVAKLNIIF
FRELNYKNSSEPSVMTVLLNSLNLAGSOWSLWFGSSVLSVMEAEILIDLLVITFLMLL
RRFRSRYWSPGRRGGTQEVASTPAASLSPSSCPHPAFSSPPDPALSPALSAPPPA
YATLGPHPAPSLAEASTSAHAPGP"
BASE COUNT 653 a 1239 c 896 g 798 t
ORIGIN

Query Match		45.0%;	Score 1401.8;	DB 3;	Length 3586;
Best Local Similarity		72.4%;	Pred. No. 0;		
Matches 2035;	Conservative	0;	Mismatches 657;	Indels 119;	Gaps 12;
QY	199	cagcccatgcaagactggggaaagggggaacaaacgtgaagagcagggcctgggcccggaa	258		
Db	691	CTGCACCTGTCAGGGTCTATGAAGGAGACAAAGCCTGAGGAGCGGGCGGCTCCGAG	750		
QY	259	ccctcagcaaccggcagcccaacggaggagagcactgattgaattccaccgctcc	318		
Db	751	CCWTCGGGCCCCCGCCGCGCACGGAGGAGGAGGCCCTGCTCGAGATTCCACCGCTCC	810		
QY	319	taccggagctcttcagttctctgcaacaacaccacatccacgggcccattccgctg	378		
Db	811	TACGGGAGCTCTTCGAGTCTTCTGCAATAACACCACCATTCCAGCGGCCATCCGGCTG	870		
QY	379	gtgtgtcccaaacacacacccatgaagacggcctcttggcgggtgctgtgctgtgcaac	438		
Db	871	GTGTGCTCGCAGCACAAATCGCATGAGACAGTCTTCTGGGCCGTGCTCTGGCTCTGCACC	930		
QY	439	ttcggcatgattactggcagtttcgccttgcctgtgttcgagaggtacctcagctaccagty	498		
Db	931	TTCCGGATGATGTACTGGCAGTTTCGGCAGCTCTTCGGAGAGTACTTCAGCTACCCCGTC	990		
QY	499	agctcaacatcaacctcaattcagacaagctggtttccctcctcctcactgctgcaac	558		
Db	991	AGCTCAACATCAACCTCAACTCGGACAAGCTTGTCTTCCTCGGGTCTCCATCTGGACC	1050		
QY	559	cttaactcttacagatacactgaaattaaagagagctggaagagctggaccgcatacag	618		
Db	1051	CTAAACCCCTFACAGGTACAAGGAAATTCAGAGGAGCTGGAGAGCTGGACCGCATCAG	1110		
QY	619	gagcagacgctttttgacttgtaacaatacaactcttcctacactgcgcaggctggggcc	678		
Db	1111	GAGCAGACTCTTTTCGACCTGTACAAGTACAACCTCTCCAAGACCCCTC-----GTGGCC	1164		
QY	679	cgacgcgagctccgcgacctctctgggtgctttccgcacccccctgacgcctgacgc	738		
Db	1165	CAGCGCGCTCCCGCGGACCTTGGGGAGCCGCTGCCGACCCCTTACGCGCTGCC	1224		
QY	739	actccacctccgcctactccgcgcagcgcgagcggggtctctccagcgtacgcgac	798		
Db	1225	GTCCCGGCTCCGCCACGCGGCCCGGGAGTCCCGCGCGCGGCTCCACGATCGGGAC	1284		
QY	799	aacaaatcccaagtgaacgcggaaggactggaagatcggtcttccaaactgtgcaaccgaac	858		
Db	1285	AACAACCCCGAGGTGAACAGGAAGGACTTGAAGATCGGCTTCCAGCTGTGCAACCAAGAC	1344		
QY	859	aaatcagactgtttctacagacatactcctctgggttgatgacagtgaaggagtgtac	918		
Db	1345	AAATCGAGCTGCTTCTACACAGAGCTGCTCAGGGGTGTGCTGTGAGAGAGTGGTAC	1404		
QY	919	cgcttccattacatcaactctgtccagactgtc---ggacacctgcgcgcgtctctag	975		
Db	1405	CGCTTCCACTTACATCAACATCTCTCCAGACGGAGACAGACACCTCTCCCTCGCTGGAG	1464		
QY	976	gaagaagccctgggaacttcattcttccactgtctgcttcaaccaggccccctgcaaccag	1035		
Db	1465	GAGGAGCTGTGTGGCAAAATTCATCTCACCTGCGCGCTTCAACCAAGGACTCCTGCAATGAG	1524		
QY	1036	gcgaattattccaagtctccaccacccatgtacggaactgctacactttcaatgacaag	1095		
Db	1525	GCGAATTTACTCTCATTTCCACACCCGATGTATGGAAATCTGTACACTTTTCAACGACAG	1584		
QY	1096	aacaactccaatctctggtctcctccatgcctggagtcacaatggttctccttgaca	1155		

Db	1585	AACAGCTCCAACCTCTGGATGTCTCCATGCCGGGTCAACAACGGTCTGTCTCCTGCACA	1644
QY	1156	ctgcacagagcagagaactgaactcaatccccctgctgtccacagtgacgggggcagagtg	1215
Db	1645	CTGCGCACAGACAAAATGACTTCAATCCCATCTGCTGTCTCCACGGTGTGACGGGGCGGGTG	1704
QY	1216	atggtgcattggtcaggtgagcctgctttatgattgattggtgcttcaacttgagcct	1275
Db	1705	ATGGTGCATGAACGGGATGAGCCTGCTTTATGGATGATGCTGGCTTTAATCTCGGGCCT	1764
QY	1276	ggcgtggagacctccactcagtatgagaaaggccctggacagcctctcggaggaaattac	1335
Db	1765	GGCGTGGAGACCTCCCATCAGCATGAGTAAGSAAGCTGTGGACAGACTTGGGGTGACTAT	1824
QY	1336	ggcagctgtactgagaatgtagcagatgtcccggfcaagaacctttacccttccaaagtat	1395
Db	1825	GGGCACTGCAACCAAGAAATGGCAGGAGTCCAGTGGAGAACCTTTACAATACCAAGTAC	1884
QY	1396	acacagcaggtgtgcattcactcctctccagagagacatgatcaagaagtgtggtctg	1455
Db	1885	ACGACGAGGTGTGCATCCATTTCTGTGTTCCAGGAGAGCATGATCAAGGAGTGTGGCTGT	1944
QY	1456	gcctacattcttaccctaaagccaaggagttgagttctgtgactaccccaagacagagc	1515
Db	1945	GCCTACATCTTCTACCGCGCGCGACGGCGTAGAGTTCTGTGACTACAGSAAGCATAAAT	2004
QY	1516	tccctgggctattgctattataaactgcagggcgctctcctctggacagcctgggctgt	1575
Db	2005	TCCTGGGGCTACTGCTACTATAAGCTCCAGGATGCCCTTCTCCTCAGACCGCCTGGGTGT	2064
QY	1576	ttctcaagtgtcggaagcctttagtgtatcaactacaaactcctgcgcggtactca	1635
Db	2065	TTCAACCAAGTGCCEGAAGCCATGCAGTGTGACCATCTACAAAGCTCTCTGCAGTTACTCA	2124
QY	1636	cgggtggccactctgtgaagtccaggatttgatcttcgagatgctgtccttgcagaacaat	1695
Db	2125	CAATGGCGCTCCGGACATCCAGGACTGGGTCTTCCAGATGCTTCCCGACAGACAAC	2184
QY	1696	tacactattaacaacaaaaaacggaggtgcaaaagctcaacatctcttccaaaggactg	1755
Db	2185	TACACCATCAAGAACAAAAGGGATGGAGTTGCCAAATCAACATCTTCTTCAAGGAGCTG	2244
QY	1756	aactataaaactaattcggagctcctctgtcacgagtggtcacgcctcctgtccaaactg	1815
Db	2245	AACTACAATCTAATCTGTAGTCTCCCTCAGTACAGATGGTACAGCTCTCTTCCAACTGT	2304
QY	1816	ggcagccagtgagcctgtgtgttggtcgtccggtcctcctgtgtgtggagatggcgagac	1875
Db	2305	GGCAGCCAGTGGAGCCTGTGTTCCGGCTCTCCGTGCTTCCGTGCTGGAGATGGCTGAG	2364
QY	1876	gtcatcttcgacctcctgtgtcaatacaacttctcatgtgtacgcggttccggagcgg	1935
Db	2365	CTCATCTTACCTCTCTGTTATACCTTCTCATGTCTTCCCGAGGTTCCTCGAAGCCGA	2424
QY	1936	tactgtctccagacagagggccaggggtgccagagagtgccctccactcccagcttcc	1995
Db	2425	TACTGTGTCACGCGGAGGGGCAAGGCAACCCAGGAGGTGGCTCCACTTCCAGCGGCC	2484
QY	1996	tccttcccgctccggtttctgtcctcaccttaca tccccacacaccttcttggccccagag	2055
Db	2485	TCCTTCCCTCCAGTTTCTGCCCCACCCAGCTTCTTCTCTCATCCCGCGCACACCT	2544
QY	2056	ggatgacctcctccctggccttgacagccctccactgcctatgtactctagggccc	2115
Db	2545	GCCATCTCCCC-----GGCCTTGTCAAGCCCTCCACTGCCTATGCCACCTGGGCCC	2598
QY	2116	agtgccccctccactggaactcctggcgccctgactgttctcctgtgccctggcggtc	2175
Db	2599	CACCCAGCTCCATCAGGCTTGGCAGAGCCGACACTCTGCCCCAGCTTCCGGGGGAGGCC	2658
QY	2176	tgagagggagagatctctctcaccagcctgagctccctt-----	2218

Db	2659	TGAGGGAAGGACGGTATCCCTCCCAAGGACGACATTCCTTGGTGGGAGGACCGAG	2718
Qy	2219	--gtaaactgatgatctactcttcagacacatcttccccagcgcgcgcgcgcgcgcgc	2276
Db	2719	CCTTGGCAGGATCAAGAATGTTTGGCGCTTCCCTGTCCAGAGCTGCCACGCTGCCCTGG	2778
Qy	2277	tgtg-----gtgctgctgctgaggaacagcgtgagtaaggggccacgaagaagcgtcca	2331
Db	2779	TGTGTGCTGGTGGCGGGGAGATGAAGCAGCATGGATAGGGGCTGTGAAATTTGCCTG	2838
Qy	2332	gagagcaggggctaacgaactctctcagagctgcctcgcctccctgc-----	2376
Db	2839	GAGACCCCGACCGAGG-ATCACCCAGTTCCTCCCGTACCCTGGTCTTCCTCGCAGAG	2897
Qy	2377	--tctgaacactgccttcacaaagcacagacaagtctcctcttccctcgcgactcagccag	2434
Db	2898	ACACTTTTGCTCTCCCCACCCAGTCAGCGCAAGTCCCTCTTTTCTTGGATCGGCTGAG	2957
Qy	2435	ccagaactggagctctgacaaagaaacttctctgggagaggaccagaacaaatagaaca	2494
Db	2958	CCAAACTTGGATCTCTGACAAACAACCTTCTCTAGGAAACAACCTGACGACGACACAAAC	3017
Qy	2495	gccatgcactctacccccgtga-----	2520
Db	3018	ACAGACGAGGGCACACAAGGCGTGTGGTTCTTCTTAAAGATGGCTGGATTGGGCC	3077
Qy	2521	agaggctggcctggcctcactgccttcaggagacactgatgctcgtcctctccttgaact	2580
Db	3078	CCAGAATGGCCCTCCACACTGCCTTCCAGTGACACAGCGCTTTGGGCTCTCTTTGAAC	3137
Qy	2581	tgggtggggaaccccaaccccaagccctt--gtagtcttttgggcaattccccctccc	2638
Db	3138	GGGGTGGGAAGCCACCCCAAGCCGCCCTCGGTTAGTTCTGCTAGGCGACTCTGTTTGCC	3197
Qy	2639	tcaactcctcagagtgaggactagataagacagaccagggtagccctgggtgtctctctc	2698
Db	3198	TGACTCC-CAGGGCAGGAACCTGGGCGACCGCGATGTATTAAAGACCTGGATTCCAGTCA	3256
Qy	2699	attctcaagactctctcattgtatcacacataccccagctgcctctgtggtgtgtctcct	2758
Db	3257	GCCTCACTACCCCATCCCTCTGACGGAGCTCGCCGGCGCTCTGCCCTGTGTCCTTGT	3316
Qy	2759	tcttt-----ctgttctcagcctgggaagctcttcttcactg	2795
Db	3317	CCTTTGTAACCTCTGGCTACTTGTCTCAAGCTATTTCCTTAGCTGGACGCTCCCGGCC	3376
Qy	2796	tccctcagaacacacagctatgatcctctagaaacctgccaaaagggcaccatttcttt	2855
Db	3377	CCATCTCTGGGAGAACTCCTATTGCATCCCTGAGAACCCGTGTC--GGATACCACTTACTTC	3434
Qy	2856	tgtgaactcagttaccctgttatgtctccccagaattgtcgccctctccc	2906
Db	3435	TGTGAGGGCTTCTTCCCAACCTCTTCTCCCGAGATCTTACCACTTCCCTC	3485
RESULT	9		
LOCUS	CP0249296		
DEFINITION	Cavia porcellus mRNA for sodium channel, alpha subunit.	ROD	29-MAR-2000
ACCESSION	AJ249296		
VERSION	AJ249296.1	GI:6687302	
KEYWORDS	alpha subunit; sodium channel.		
SOURCE	domestic guinea pig.		
ORGANISM	Cavia porcellus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.		
AUTHORS	Schneider M., Mastroberardino, L., Reifarth, F., Weber, W.M., Verrey, F. and Claus, W.		
TITLE	CAMP sensitivity conferred to the epithelial Na+ channel by alpha-subunit cloned from guinea-pig colon		
JOURNAL	Pflügers Arch.	439 (5),	579-587 (2000)
MEDLINE	20225288		

[illegible]


```
Db 1426 TCCTGTTACTCGCATGGCCCTCGGTGACGCTCCAGGAGTGGGTCTTCCAGATGCTGTC 1485
Qy 1683 cttgcagacaattacactattataacaaagaagaacaggagttgcaaaagctcaacatctt 1742
Db 1486 GCTACAGAACAACTACACGGGTGAGCAACAAGAGAATGGGTGCAAGCTCAACATCTA 1545
Qy 1743 cttcaaggagctgaactataaactaaatcggagctctcttctgtcacgagtggtcagct 1802
Db 1546 CTTCAAGAGTTGAACACTACAAGCCAAATCTGAGTCCCTCTGTGCAGATGGTCACCT 1605
Qy 1803 cctgtccaaactgggagcagctggagcgtgtggttttgctgtcgtctgtctgtgtgt 1862
Db 1606 COTGTCCAACCTGGGCAGCAGCTGAGGCTGTGGTTCGGCTCTTCCGTCTCCGTGGT 1665
Qy 1863 ggagatcgggagcgtcaattcttgagacctcctggtacatacacactctcactgctgcgcg 1922
Db 1666 GGAGATGCGCGAGCTCCCTTCGACCTCTCAAGTCATCATCTTCCTCATGCTGCTCCGGCG 1725
Qy 1923 gttccggagcgggtactggtctccaggagcagggccaggggtgccaggaggtggcctc 1982
Db 1726 GTTCCGACCGCTACTGCTCTCTGGCGGGGCGCGGGGCGGGAGGTGGCCTC 1785
Qy 1983 cactcagacttctccttcccgctccgctgttctgtcctcactacataccccaccacttc 2042
Db 1786 CTCGCGGTGTCCGGCTGCCCTCCGCTTCTGCCCCCAGCCACGCTCC-----CCGTC 1839
Qy 2043 ttgtccccagcagggcatgacccctccctggcctgacagccctcactcactgctatgc 2102
Db 1840 CGTGGCCCCAGCGTGGCCCCACACTGCCCCCGCTCTTGACAGCCCCCTCCACCTGCGCTATGC 1899
Qy 2103 tacttagccccagtgccccctcactgactcgtcggcgctgactgtctgtcctgtgc 2162
Db 1900 CACTTTCGCCCTCCCTGCTCATGCTCGGC-----TCGCGCTGTGC 1941
Qy 2163 cctgctggcgtctctgagagaggaggaatcctctcaccagcctgactgactcctgttaa 2222
Db 1942 CCTGGGGAGCCCTAAGAGGAGACAGTGCCTCCGCCGCCGCCCGCGGTG----CCAGGGGAA 1997
Qy 2223 actgatgatattcacttcagacacattctcccgagcgccccgcgagcctctctgtgt 2282
Db 1998 GCCTTGGCAGAATGAAGATGTTTGGGGCTTTCCTCTAGAGTGGCCCACTGCCCATGGT 2057
Qy 2283 gtgcctgctgagggaaacagcgtgtagtaaggggccaggaagccgtccagagagcaggg 2342
Db 2058 GGGC-----GGGGAAGGGGCCCCCGGAAGTTGTCCAGAGAGCAGCGG 2098
Qy 2343 ctaacga-tctgctcagagctgcccctgcccctgtcttgaaacactgcttcacacagcac 2401
Db 2099 CCGGTGAGGCTGCCAGAACTGCCCTGTGCTGTGCCGGTTCCTCTGGTCTCC--CCTG 2156
Qy 2402 agacaagttcctcttccctcggatcagcccaagcagacttgagctctgacaaagaaact 2461
Db 2157 AACACTGCGCTCCCTCCCAAGTGCAGGCGCGAGCCAGACTGGGAATGCTGCAAGGAAC 2216
Qy 2462 ttcctggagagaccaggg-----aacaataagacagccatcgactcctaccgccgtg 2517
Db 2217 TTTCTAGGAGACGACCAACAGCAGCAACACACAAACAAAGCCCAAGGGCTGCACAGGTC 2276
Qy 2518 acc-----agaggtggcctggcctcactgccttccctcagggacactgat 2559
Db 2277 TCCTACGCAATGGTGACTGATGGCCAGAGCCAGCCCGCCGCTGGCTGGGGCCACAGAT 2336
Qy 2560 gtctactcctcttctgacttggttggggaacccccccccccccccccc-ttgttagtt 2618
Db 2337 GTCTGCTGTGCTTGAACCTGGAGTGGGAACCCCAACCCCAAGGCCCTTTATTCGTT 2396
Qy 2619 ctttggaattccctccctcactcctcactcaggttggtgggactaggttaagacagcaggg 2678
Db 2397 CTTGGCAATTCCTCTTCCCTGGCTTGCAGGGTGAAGGCTGAGCAGAGCGCTGTGTGTA 2456
Qy 2679 taggctgg 2687
Db 2457 CAGGCGGG 2465
```


Qy	289	gaggaggcactgattgaattccacgcgcctcctaccggggagcctctctccagttctcttgcac	348
Db	675	GAGGAGCCCTGATAGAGTTTCCACCGCTCTTACCOCGAGCTCTTCCAGTTCTTCTGCAAC	734
Qy	349	aaacaccattccacggggccatccgcctgggtgctgtccaaacacacgcgatagaagc	408
Db	735	AACACCACATCCATGGCGCCATCGCCTCTGCTCTCAAGCACACACCGGATGAAGACG	794
Qy	409	gccttctggggcgtgctgtggctgtgcaccttcgcgcatgagtactggcagttgccttg	468
Db	795	GCUTTCTGGCCGCTGCTCTGGCTCTGTACATTTCGGCATGATGTACTGSCAGTTGCGCTTG	854
Qy	469	ctgttcgaggagtaactcagctaccagtgagcctcaacatacaacctcaatcgaacaag	528
Db	855	TTGTTTCGGAGAGTACTAGTTTACCCGGTCAAGCTTCAACATCAACCTCAACTCCGACAG	914
Qy	529	ctggtctccctccgtcactgctctgcaccttaactcttacagatacaactgaaattaaa	588
Db	915	CTGGTCTTCCCTCGGGTTACTGTCTGCACCTCAATCCTTACAGGTACAGGAAATTTAA	974
Qy	589	gaggagctggaagagctgacgcgcatcacgggagcagacgcttttgacttgtacaatat	648
Db	975	GAGCAGCTGAGGGAGCTGACCGCATCACACAGCAGACGCTCTTTCGACCTGTACAATTAC	1034
Qy	649	aactcttctctaactgcacagcctggggccgcgacgcgcgcagctccgcgacctcctgg	708
Db	1035	AACGCTTCCAGCACCTCTCTGGCCGGGGCTCGACACCGCGCGCACGCTC-----	1082
Qy	709	gctttccgcgaacccctgcagcgctcggcactccacactccgcctactccgcgcgcag	768
Db	1083	GCAGACACCTTGCCTGTACCGCTGCAGCGCATCTCCCGTGCAGCCGACCGCCCGCAGAGCG	1142
Qy	769	gcgcgacggcgggtcttctccagctacgcgacaaataccccaaagtggacggagagactg	828
Db	1143	CGCAGCAGCAGACCCCTCCAGCGTGCAGGACACAAACCCCGCTGGACCGGAGAGACTGG	1202
Qy	829	agaatcggttccaactgtgcacaacagaacaaatacagactgttttaccagacatactcc	888
Db	1203	CGGTCGGCTTCCAGCTGTGTAAACAGAAATAATCGGACTGCTTCTTACGACAGCTCCTCG	1262
Qy	889	tctgggttgatcagtgaggagtggtaccgcttccattacatacaactctgtctccaga	948
Db	1263	TCAGGGGTGACGGAGTGGCGGAGTGGTACCCTTCCACTATCATCAACATCCTGGCCCAA	1322
Qy	949	ctgtcggcacacctcgccgctctagagggaagccctgggcacactctcatcttcaactgt	1008
Db	1323	GTGCTGACACCTCACCTCGCTGGGAGGAAGAACCTTGGGCAACTTCATCTTCGCCTGC	1382
Qy	1009	cgcttaaccaggccccctgcacacaggcgaattattccaagtttccacccccatgtac	1068
Db	1383	CGCTTCAACCCAGCCCGCTGCACCCAGGAGAAATTACTCTCACTTCCACACCCAAATTTAT	1442
Qy	1069	ggaaactgtcaacttcaatcaacaagaacaaactccaatctctgagtgctccatgcct	1128
Db	1443	GGCAACTGTCTACACTTTCAACAAACAAAGAATGACTCCAGCCTCTGGATGGCTTCCATGCCT	1502
Qy	1129	ggagtcaacaatggtttgctcctgcacactgcgcacagagcagaatgacttcatccctcg	1188
Db	1503	GGGTCNAACAANTGGTCTGCTCCCTGACGCTGCGCAGCGAGCAGATGACTACATCCCCCTG	1562
Qy	1189	ctgtccacagtgacggggccagggtgtggtgcattgtgcaggatgagcctgaccttatg	1248
Db	1563	CTGTCCACGCTGACGGGGCCAGGGTTAACAGTGCATGGCAGGATGAGCCGGCTTTTATG	1622
Qy	1249	gatgatgtgcttcaacttgaggcctggcgtgcagaacctccatacgtatgagaaagaa	1308
Db	1623	GATGATGGTGGTTTCAACTTGGCGCCTGGTGTGGAGACCTCCATCAGCATGCGGAGGAA	1682
Qy	1309	gccttgacagcctcggaggaataacggcactgtactgagaatgtagcgatgtcccg	1368
Db	1683	GCCCTGGACACAGCTTGGGGCAGTATATGGAAGACTCCACCCAGGACGGCAGCAGCTCCCT	1742
Qy	1369	gtcaagaacctttacccttctccaagtatacacagcagggtgtgcttaccctcctgctccag	1428

[illegible]

Db 736 CAGACGTACTCTCAGGGGTGGATGAGTCGCGGAATGTTACAGCTTTTCACTACATCAAT 795
Qy 937 attctgtccagactgtcgcagacacctccgcgtctctagagaaagaacccctgggcaacttc 996
Db 796 ATCTCGGCACAGATGCTGATGCC---AAAGACCTGGATGAGTCTGACTTTTGAGAAATTT 852
Qy 997 atcttcacactgtcgtctcaacaggcccccctgcaacaggcggaattattccaagtctccac 1056
Db 853 ATTATGCTTTGCCGCTTCAATGAAGCAACATGTGACAAAGCGGAATTACACCCACTTCCAC 912
Qy 1057 caccacctagcgggaactgtacactttcaatgacaagaacaactccaatctctggatg 1116
Db 913 CATCCCTTGTATGGGAACGTCTACACTTCAATGACAATAGCAGC---AGCCTGTGGACA 969
Qy 1117 tcttcacactgtggagtcaacaatggtttgtccctgacactgcgcacaagcagaatgac 1176
Db 970 TCTCTCTGCCCTGGGATCAATATGCGCTCTCTCGTGGTGGCCACCGAGCAGATGAT 1029
Qy 1177 ttcatcccccctgtctccacagtgcagggggccagggtgatggtcatggtcaggatgag 1236
Db 1030 TTTATCCCTCTGTTCAACGGTAAGTGGAGCCAGGGTCATGCTCCACCATCAGATGAG 1089
Qy 1237 cctgcctttatgatggtggtctcaacttgagcctggcgtggagacctccatcagt 1296
Db 1090 CCAGCCTTTCATGGATGATGGGGGTTTCAATGTGCGTCCAGGTATTGAGACCTCCATCAGC 1149
Qy 1297 atgagaagaagccctgacagcctcggagagaaatacggggactgtactgagaatggt 1356
Db 1150 ATGAGGAAGGATGACTGAGCGCTCTGGGGGAGTTTACAGCGACTGTACAGAGGATGCC 1209
Qy 1357 agcgatgtcccggtcaagaacctttacccttccaagtatacacagcaggtgtgtcattcac 1416
Db 1210 AGCGATGTCCAGTGCAAAATCTGTACTCATCTCGTACATCAACAGGTCTGCATTCGT 1269
Qy 1417 tctcgtctccaggagaacatgatcaagaagtgtggtgtgcctacatactttaccctaag 1476
Db 1270 TCTGTCTTTTTCAGCTCAACATGTTAAAGCGCTGCTCTGCTGCTATTTATCCCTTAA 1329
Qy 1477 cccaaggagattgattctgtactaccgaagaagcagagctcctgggctatgctattat 1536
Db 1330 CTGTGATGGAGCAGATGATTGTGACTACACAAAGCAGTGTGGGGCTACTGCTATTAC 1389
Qy 1537 aaactgcaggcgcccttccctggacagcctgggctgttctcccaagtgtcggaaagcct 1596
Db 1390 AAACCTCTTGTGAATTCAAAGCTGATGTGCTGGGCTGCTTCCACAAATGTCCGAAACCT 1449
Qy 1597 tgtagtgtgatcaactacaactctctgcggctactcacggtggccatctgtgaagtc 1656
Db 1450 TGCAAAATGACAGAAATACCAGCTCTCAGCTGGATACTCTCGCTGGCCTTCTGCTGTCTCA 1509
Qy 1657 caggattgattctcgagatgctgtccttgagaaacaatacaactatcaacaacaaga 1716
Db 1510 GAGGATTTGGGTTTTTACATGCTTTTACACACAGAAACAAATACAAATCAATCAAGAGG 1569
Qy 1717 aacggagttgcaagctcaactcttctcaaggagctgaactataaaactaattcggag 1776
Db 1570 AATGGAGTTTGCAAAAGTGAATATCTTTTGGAGGAATGAACCTACAAAGCAATGGGGAG 1629
Qy 1777 tctccttctgtcacagatggtcagcctcctgtccaacctgggcagccagtggaacctgtgg 1836
Db 1630 TCACCAGCCTTTACAGTAGTACCCCTCTTGTCTCAGCTTTGGGAATCAGTGGAGTCTCTGG 1689
Qy 1837 ttggtcgtccgtctctctgtggtgagatggtggagcgtcatcttcgacctcctggtc 1896
Db 1690 TTTGGATCTCTGCTCTATCAGTGTGGAGCTTGCAGAGCTGATTTCTGGATTTTCACTGTC 1749
Qy 1897 atcacacttctcatgctgactacgcgggttccgcgagccggtactgg 1941
Db 1750 ATCACTTTTCACTGCGCTTCGCTGGTTCGCTCCAGCGGTGG 1794

RESULT 13

AF232715 1003 bp mRNA MAM 01-MAR-2000
LOCUS Ovis aries amiloride-sensitive sodium channel alpha subunit mRNA,
DEFINITION partial cds.
ACCESSION AF232715
VERSION AF232715.1 GI:7110531
SOURCE sheep.
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Caprinae; Ovis.
REFERENCE 1 (bases 1 to 1003)
AUTHORS Barley,J., Syrris,P. and Walters,D.
TITLE Expression of sheep lung amiloride-sensitive epithelial sodium
channel alpha subunit gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1003)
AUTHORS Barley,J., Syrris,P. and Walters,D.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2000) Child Health, St. Georges Hospital Medical
School, Cranmer Terrace, Tooting, London SW17 0RE, UK
FEATURES
source
1..1003
/organism="Ovis aries"
/db_xref="taxon:9940"
/tissue_type="lung epithelium"
CDS
<1..>1003
/note="OENac"
/codon_start=1
/product="amiloride-sensitive sodium channel alpha
subunit"
/protein_id="AAF36982.1"
/db_xref="GI:7110532"
/translation="KDWKGFQLCNQKSCFYQYVSSGVDAVREWRPHYINILSR
RODTSPLEEDVLGRFIETCRNQDSCNARYSHFHPMYGNCYTTFNDKSNLWISS
VPGVNSLSTLRTEQNDIFPLLSVTGARVMVHDEPAFMDQGFNLRPGVKTSIS
MSKEAVDLRGDYGDCITKNGSEVPVGNLYTTQOVCIHSCFOESMIKECGCAIFY
PREGVEFCDDYKKHNSWGVCYYKLQDAFSSDRLGCTFKCRKPCSVTYLKLSASYOMP
SVTSQDWVFQMLSRQNNYTIKNKRDGVAKLNITFFKELNYKSNSSEPSVTMTLLSNLG
"

Query Match 24.1%; Score 751.8; DB 3; Length 1003;
Best Local Similarity 85.0%; Pred. No. 1.7e-170;
Matches 853; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

Qy 820 aaggactggaagatcgcttccaactgtcaaccagacaacaaatcagactgtttctaccag 879
Db 1 AAGGACTGGAAGATCGCTTCCAGCTGTGCAACCAAGAAATCGGACTGCTTCTACCAG 60
Qy 880 acatactcctctgggtggatgcagtgagggagtggtaccgcttccattacatacaact 939
Db 61 ACATATCGTCAGGGGTGGATGCTGTGAGGAGTGTGTATCGCTTCCACTATATCAACATT 120
Qy 940 cgtgtccagactg---tcggacaacctgcgcgtctctagagaaagaacctcgggaacttc 996
Db 121 CTGTCCAGAGGAGACAGGACACCTCTCCCTCGCTGGAGGAGGACGTGCTGGGCAAAATC 180
Qy 997 attctcacctgtcgtctcaaccaggccccctgcaaccaggcgaattattccaagtctcac 1056
Db 181 ATCTTCACTGCGCGTTTCAACACGAGACTCCTGCAATGAGCGAATTTACTCTCACTCCAC 240
Qy 1057 caccctcatgacgggaactgctacactttcaatgacaagaacaactccaactcctcatgag 1116
Db 241 CACCCGATGTACGGAACACTGCTACACTTTTCAACGACAAAGAACAGCTTCCAACCTCTGGATA 300
Qy 1117 tcttccatgcctgagtcacaacatggtttgtcctctgacactgcgcagcagcagaatgac 1176
Db 301 TCCTCCGTGCCGGGTCAACACGGCCTGTCCCTGCACACTGCGCAGACAGACAAATGAC 360

QY	1177	ttatccccctgctgtccacagtcagcggggccagggtgatggtgcatggtcaggatgag	1236
Db	361	TTTCATCCCACTGCTGTCCACGGTGACCGGGGCCGGGTGATGTCATGAACGGGATGAG	420
QY	1237	ctgctcttatgatgatggtggtctcaacttgaagccttgagcctggtgagacctccatcagt	1296
Db	421	CTGTCCCTTTATGATGATGCTGGCTTTAATCTGCGGCCGGCGGTGAAGACCTCCATCAGC	480
QY	1297	atgagaagaaacccctggacacccctcgagagaaataacggcgactgtactgagaatggt	1356
Db	481	ATGAGTAAGGAACCGCTGGACAGACTTGGGGGTGACTATGGCGACTGCACCAAGAATGGC	540
QY	1357	agcagatgtcccggtcaagaacctttacccttccaagtatacacagcaggtgtgcaattcac	1416
Db	541	AGCGAGTCCCGGTGGGAACCTTTTATAATACCACGTACACGACGACAGGTGTGCATCCAT	600
QY	1417	tctgtctccagagagaacatgatcaagaagtgtggtgtgcttcacatctcttaccctaaag	1476
Db	601	TCGTGCTTCCAGGAGAGCATGATCAAGGAGTGTGGCTGTGCTTACATCTTCTTACCCACGG	660
QY	1477	ccaaagggagttagttctgtgactaccgaaagacagagctctctgggctattgctattat	1536
Db	661	CTGAGGGCGTAGATGCTGTGACTACAGAAGCATAAATCTTTGGGGCTACTGCTACTAT	720
QY	1537	aaactgcaggcgctctctctgtggacagcctgggtgttcttccaaagtgtcgaagcct	1596
Db	721	AAGCTCCAGGATGCCCTTTTCCTCAGACCGCTGGGCTGTTTCACCAAGTGC	780
QY	1597	tgtagtgtatcaatacaaaactctctcggtctactcaogtgcagtggtccatctgtgaagtc	1656
Db	781	TGCAGTGTGACCATCTACAAAGCTCTCTGCCAGTTACTCAGATGGGCCCTCCGTGACATCC	840
QY	1657	caggattgatcttcagatgtgctctgtcagaaacattacactattaaacacaaaga	1716
Db	841	CAGGACTGGGTCTTCAGATGCTGTCCCGGCAGAACTACACCATCAAGAAACAAAGG	900
QY	1717	aacggagtgcgaagctcaacatctcttcaaggagctgaactataaaactaatctcgag	1776
Db	901	GATGGAGTTGCCAACTCAACATCTTCTCAAGGAGCTGAACCTACAAATCTAATCTCTGAG	960
QY	1777	tctctctgtcagatggtgcagcctctgtccaaacctgggca	1819
Db	961	TCCTCCCTCAGTCACGATGGTCAACCCCTCTGTGCCAACCTGGGCA	1003
RESULT 14			
GGU62902			
LOCUS	2321 bp	mRNA	VRT 27-MAR-1997
DEFINITION	Gallus gallus amiloride-sensitive epithelial sodium channel alpha subunit (alpha cEnAc) mRNA, complete cds.		
ACCESSION	U62902		
VERSION	U62902.1 GI:1497998		
KEYWORDS	chicken.		
SOURCE	Gallus gallus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus		
REFERENCE	1 (bases 1 to 2321)		
AUTHORS	Killick,R. and Richardson,G.		
TITLE	Isolation of chicken alpha ENac splice variants from a cochlear cDNA library		
JOURNAL	Biochim. Biophys. Acta 1350 (1), 33-37 (1997)		
MEDLINE	97157073		
REFERENCE	2 (bases 1 to 2321)		
AUTHORS	Killick,R. and Richardson,G.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-JUL-1996) School of Biology, Sussex University, Falmer, Sussex BN1 9QG, UK		
FEATURES	Location/Qualifiers		
source	1..2321		
	/organism="Gallus gallus"		
	/strain="Isa Brown"		
	/db_xref="taxon:9031"		

gene	/tissue_type="cochlea"				
	/clone="2.5kb transcript"				
	89..2002				
	/gene="alpha cENaC"				
CDS	89..2002				
	/gene="alpha cENaC"				
	/note="contains two hydrophobic domains, a cysteine rich domain, and a proline rich region"				
	/codon_start=1				
	/product="amiloride-sensitive epithelial sodium channel alpha subunit"				
	/protein_id="AAB50550.1"				
	/db_xref="GI:149799"				
	/translation="MGTASRGGSVKAEMPEGEKTRQCKOETEQOQKEDEREGLIIFY				
	GSYQDVQFPCSNTHGAIKRLVCSKKMKMTAFWSVLFILFTGLMWQFGLVREYF				
	STPVLNLSNDRILFPAPVLTCLNPIRYSAIRKKLDELQDHTQLDLYDNMSL				
	ARSDGSAQSHRSTRSLHHVQRHPLRQKRDNLVSLPENSVDKNRWIKGFVLC				
	ENNECDFHOTYSYSGDAVREWSFYINILAQMPDAKDLDEDFENFIYACRNEATC				
	DKANTYFHHPLFGNCYTFDNNSSLSLPGINNGLSLVTRTEQNDFTPLSTVTG				
	ARVMYHDQNEPAFMDGDFNVRPGIETISIMRKEMTERLGSSYSDCTEDGSDVPQNL				
	YSSRTEQVCIKSCFQLNMKRSCAVIYFPLPDGAEYCDYTKHWANGCYKYLAEF				
	KADVLGCFHKRKPCKMTETQLSAGYSRWPSSAVSDWYFMLSQKNYITSRKNGVA				
	KNVIFEEWNYKTNGESPATVVTLISQLQNSLWFGSSVLSMELAEILDTFVIT				
	FLIAFRKRSQKHSSPAPPNSHDNTAFQDEASGLDAPHRTVEAVVTTLPSYNSLE				
	PCGPSKDGEGTGLE"				
BASE COUNT	584 a	584 c	578 g	575 t	
ORIGIN					
Query Match	24.0%; Score 747.4; DB 4; Length 2321;				
Best Local Similarity	66.4%; Pred. No. 2.1e-169;				
Matches 1145; Conservative	0; Mismatches 556; Indels 24; Gaps 4;				
QY	217	gggaagggggcaaacgtgaagagcagggctggggccgggaacccctcagaccccgagcag	276		
Db	125	GAGAAGATGCGCAGAAGGGGAGAACGAGCAGTGC AAGCAGGAAACGGAGCAGCAGCAG	184		
QY	277	ccacccgagagagagagcactgattgaattccacgcctctaccggagctcttcacag	336		
Db	185	AAGAAGATGAGAGGGAGGGCCTCATCGAATTCAT TATGGCTCTTACCAGGACGTGTTCCAG	244		
QY	337	ttcttctgcaacaacacacccatccacggggccatccgcctgggtgtgctcctcaaacacacac	396		
Db	245	TTCTTCTGTAGCAACACGACCATCCACGGAGCTAT TCGCCTGGTGTGCTCTATAAAGAAC	304		
QY	397	cgcataagacgcgctctctggcggtgctgtggtgtgacaccttggcctgatgactgtg	456		
Db	305	AAGATGAAGACGGCCTTCTGCTGTCTCTTCAAT TCTACCTTCGGCTTAATGTATTGG	364		
QY	457	cagttgccttgccttgcgagagtagtacctcagctcaccagtgagctcaacatcaacctc	516		
Db	365	CAGTTTGGAAATCCTCTACAGGGAGTACTTCACT TACCTGTCAACCTCAACCTCAACCTA	424		
QY	517	aattcagacaagctggttctccctgcccgtcactgtctgcaccttaactccttacagatac	576		
Db	425	AACTCGGATAGGCTGACTTTCCCTGCGAGTGACG CTGTGCACACTCAACCCATACAGGTAC	484		
QY	577	actgaaataaagagagctggaagagctggaccgcatcacggagacagacgctttttgac	636		
Db	485	AGTGCCATTTCGGAAGAGCTGGACGAACCTGGA CCAATCACATCAGACACTGTTGGAC	544		
QY	637	ttgtacaaatacaactcttctacactcgcagctgc cagctggggccgc			

```
Qy 817 cgaagagactggaagatcggtctccaaactgtgaaccagagaacaatacagactgtttctac 876
Db 707 AAGAACGACTGGAATGGCTTTGTTCTGTGCAAGTGAACCAATAGGATTTGTTCCAT 766
Qy 877 caacataactcctctgggtgagatcgaggtggtacgccttccattacatacaac 936
Db 767 CAGACGTACTTCCACAGGGTGGATGTCAGTGCAGGAATGGTACAGCTTTCACATCAAT 826
Qy 937 attctgtccagactgtcgggacacctccgcctctagaggaagaagccctgggcaacttc 996
Db 827 ATCCTGGCAGATGCTGTATGCC--AAAGACCTGGATGAGTCTGACTTTGAGAATTT 883
Qy 997 atttcaactgtgtcttaaccaggccctgcaaccaggcgaattattccaagtccac 1056
Db 884 ATTATGCTTGGCGCTTCAATGAAGCAACATGTGACAAGGCGAATTACACTCACTCCAC 943
Qy 1057 caaccaatgtacgggaactctacactttcaatgacaagaacaactccaatctctgtgat 1116
Db 944 CATCCCTGTATGGGAACTGCTACACCTTCAATGACAATAGCAGC---AGCCTGTGGACA 1000
Qy 1117 tctccaatgctgagtcacaatggtttgtccctgacactgcgcacagacagaatgac 1176
Db 1001 TCCTCTCTGGCTGGATCAATAATGGCTCTCTCTGTGTGGTGCACCGAGCAGCAATGAT 1060
Qy 1177 tcatccctctgtctccacagtgacggggccagggtgatggtgcattgtgcagatgag 1236
Db 1061 TTCATCCCTCTGTTGTCAACGGTAACGTGAGCGAGCGGTGATGGTCCACGATCAGAATGAG 1120
Qy 1237 cctgcctttatgatgatggtggtctcaacttgaagcctggcctggagacctccatcagt 1296
Db 1121 CCAGCCTTCATGGATGATGGGGTTTCAATGTGCTGTCAGGTAATGAGACCTCCATCAGC 1180
Qy 1297 atagaagaagaagccctggacagctcgaggaatattcacgcactgtactgagaatggt 1356
Db 1181 ATGAGGAGGAGATGACTGACGCTTCTGGGGGCGAGTTACAGCGACTGTACAGAGGATGGC 1240
Qy 1357 agcgatgtccgggtcaagaacctttacccttccaaagtatacacagcaggtgtgcattcac 1416
Db 1241 AGCGATGTCCAGTGCAAAATCTGTACTATCTCGCTACACTGCAACAGGTCTGCATTCTGT 1300
Qy 1417 tctgtcttcagagagaacatgatcaagaagtgtggtgtgctgactacattcttaccctaa 1476
Db 1301 TCCTGTCTTTCAGCTCAACATGGTAAGCGCTGTCTCTGTGCTATTACTTTTATCCCTTA 1360
Qy 1477 cccaagagagttaagttctgtactaccgaagacagagctcctgtggcctattgtattat 1536
Db 1361 CCGTATGGACGAGTATGTTGACTACACAAAGCAGCAGTAGCTTGGGGCTACTGTATATAC 1420
Qy 1537 aaactgcaggcgctctctctgtgacagcctgggctgtttctccaagtgtcgaagcct 1596
Db 1421 AAACCTCTTCTGTAATTCAAAGCTGATGTGCTGGGCTGTCTCCACAATGTCCGAAACCT 1480
Qy 1597 tgtagtgtcaactacaactctctgcggcttactcacggtggccactgtgaaatccc 1656
Db 1481 TGCAAAATGACAGAATACCAGCTCTCAGCTGGATACTCTCGCTGGCCTTCTGCTGTCTCA 1540
Qy 1657 caggattgattcttcagagatgctgtccttcagagaacaattacactattacaacaaga 1716
Db 1541 GAGGATTTGGGTTTTTTCATGCTTTTCAACAGCAACAAATACAATATAACATCCAAAGAG 1600
Qy 1717 aacgagtttcaagctcaacatctcttcaaggagctgaactataaaactaatctcgag 1776
Db 1601 AATGGAGTTTGCAAAAGTGAATATCTTTTGTGAGGAATGGAACTACAAGACCAATGGGGAG 1660
Qy 1777 tctccttctgtcagatggtcagcctctgttccaacctgggacccagctggagcctgtgg 1836
Db 1661 TCACCAGCCTTTACAGTAGTGACCTCTTGTCTCAGCTTGGGAATCAGTGGAGTCTCTGG 1720
Qy 1837 ttggctcgtccgtgtctctgtggtgagatggcggaagtcactcttgcacctcctgtgtc 1896
Db 1721 TTTGGATCCCTCTGCTTATCAGTATGAGGCTTGTGCTGAGCTGATTCCTGGATTTTCACTGTC 1780
Qy 1897 atcacacttctctgctgtacgcccgttccggagccggtactgg 1941
```

```
Db 1781 ATCACTTTTCATCTCGGCTTCGCTGCTTCGCTCCCAAGCAGTGG 1825
```

RESULT 15

GGU62903

LOCUS

DEFINITION

subunit (alpha cENaC) mRNA, complete cds.

ACCESSION

U62903.1

VERSION

U62903.1

KEYWORDS

SOURCE

ORGANISM

chicken.

Eukaryota; Chordata; Craniata; Vertebrata; Archosauria;

Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

1 (bases 1 to 3399)

Killick, R. and Richardson, G.

Isolation of chicken alpha ENaC splice variants from a cochlear

cDNA library

Biochim. Biophys. Acta 1350 (1), 33-37 (1997)

97157073

2 (bases 1 to 3399)

Killick, R. and Richardson, G.

Direct Submission

Submitted (04-JUL-1996) School of Biology, Sussex University,

Falmer, Sussex BN1 9QG, UK

Location/Qualifiers

1. .3399

/organism="Gallus gallus"

/strain="Isa Brown"

/db_xref="taxon:9031"

/tissue_type="cochlea"

/clone="3.5kb transcript"

89. .2002

/gene="alpha cENaC"

89. .2002

/gene="alpha cENaC"

/codon_start=1

/product="amiloride-sensitive epithelial sodium channel

alpha subunit"

/protein_id="AA050551.1"

/db_xref="GI:1498001"

/translation="MGTASRGSVKAEKMPGEKTRCKOETQOKEDEBREGLEIEY

GSVDVDFQFCSTNTTHCAIRLVCSKKMKMTAFWSVLFILTEGLMWQGLIREYF

SYPVNLNLNLSRLTFPAVLCTLNPIRYSAIRKKDELDDQITHQLDDIDYNMSL

ARSDGSAQFSRRRTSRLHHVQRHLRQKRDNLVSLPENSVDKNDKMGFVLC

ENNEDECFOHTYSSGVDAVREWYSFYINILAQMPDAKDLSDSFENFIYACRFNEATC

DKANYTHFHHPLYGNCYTFNDSSSLWTSLSLPGINNGLSLVVRTEONDFIPLSTVTC

ARYMVHDONEPAFMDGGFNVRPGIETISIMRKEMTERLGSSYSDCTEDGSDVPQNL

YSRYTEQVCIRSCFQLNMWVKRCSAIYFPLPDGAEYCDYTHKAVNGYCYKLLAEF

KADVLGCFHKRCPCMKTEYQLSAGYSRWPSVSDSEWVFYMLSQONKINITSKRNGVA

KVNIFFEWMYKTNGESPAFTVTTLLSQLNQWSLWFGSSVFLSYMELAEILIDFTVIT

FILAFWRFSKQWHSPPPPNSHDNTAFQDEASGLDAPHRTVEAVVTTLPYSNLSLE

PCGPSKDGGETGLE"

BASE COUNT 892 a 826 c 796 g 885 t

ORIGIN

Query Match

Best Local Similarity 24.0%; Score 747.4; DB 4; Length 3399;

Matches 1145; Conservative 0; Mismatches 556; Indels 24; Gaps 4;

Qy 217 ggggaagggggaacacgtgaagagcaggcctggcccggaacctcagcaccggcgag 276

Db 125 GAGAAGATGCCAGAAGGGGGAAGACGAGCAGTGCAAGCAGGAAACGAGCAGCAGCAG 184

Qy 277 cccaccagagagagagagagcactgattgaattccaccgctcctaccgggagcttccag 336

Db 185 AAGGAAGATGAGAGGAGGGCGCTCATCGAATTTCTATGCTCTCTACCAGAGCTGTTCAG 244

Qy 337 ttctcttgcacaacacacacccatccacggggccatccgctcctggtgtgtcctcaaacacac 396

Db 245 TTCTTCTGTAGCACACAGCACCATCCACGGAGCTATTTCGCCCTGGTGTCTTAAAAAGAAC 304
Qy 397 cgcataagacgagcctcttggcggtgctgtggtgctgacaccttcgagcatgactagg 456
Db 305 AAGATGAAGACGGCCTTCTGCTCTCTCTCTCATTTCTCACCTTCGGCTTAATGTATTGG 364
Qy 457 cagttgccttgctgttgtagaggaatcactcagctaccacagtgagcctcaacatacaactc 516
Db 365 CAGTTTGGAAATCCCTACAGGAGTACTTTCAGCTACCTACCTGTCAACCTCAACCTCAACCTA 424
Qy 517 aattcaacaagctggtcttccctgcgctcactgtctgcacccttaactcttacagatac 576
Db 425 AACTCGGATAGGCTGACTTCCCTGCGAGTGACGTGCTGCACACTCAACCCCATACAGGTAC 484
Qy 577 actgaataaaggagctggaagagctgacgcgcacacacacacacacacacacacacacac 636
Db 485 AGTCCATTCGGAGAGAGCTGAGCAACTGGACCAATCACACATCACACACTGTGGAC 544
Qy 637 ttgtacaatacaactcttctacactgcgcaggtggtggcgacgcgcgcgcgcgcgcgcgcgc 696
Db 545 CTTTATGACTATAACATGTCCCTGGCAGCAAGTGACGGGTGACGGGTGACGGGTGACGGGTGAC 604
Qy 697 gacctcctggtgcttcccgacacccctgcagcgcctgcacacacacacacacacacacacacac 756
Db 605 CGCACCTCGAGAGCCCTGCTCCATCATGTTCAGCGCCATC-----CTGCG 650
Qy 757 tccggcgacgc 816
Db 651 GGAGCGAGAGCGGGATAACTTGGT----CAGCTTGGCAGAAACAGTCCCTCGGTGGAC 706
Qy 817 cgggaagactggaagatcggtctcccaactgtgcaaccagaaacacacacacacacacacacac 876
Db 707 AAGAAGCAGCTGGAATAATGGCTTGTCTGTGAGTGAACAAATGAGGATGTTTCCAT 766
Qy 877 cagacatacctcctggtgggtgtagtgcagtgaggagtggtaccgcttccattacatacaac 936
Db 767 CAGACGTACTCTCAGGGGTGGATGCAAGTGGCGGAAATGGTACAGCTTTTCACTACATCAAT 826
Qy 937 attctgtccagactgctgcgacacactcgcgcgcctctagagaagaagcgcctgggcaacttc 996
Db 827 ATCTTGGCAGAGATGCTGATGCC---AAGACCTGGATGATGCTGACTTTGAGAAATTT 883
Qy 997 attctacactgtccttcaacacagggcccttgcacacagggcaattatttccaaagtccac 1056
Db 884 ATTTATGCTTGCCGCTTCAATGAAGCAACATGTGACAAGGGGAATTACACTCACTTCCAC 943
Qy 1057 caacccatgtacgggaactgtacactttcaatgacagaagaacacacacacacacacacacacac 1116
Db 944 CATCCCTGTATGGGAATGCTGACACCTTCAATGACAAATAGCAGC---AGCCTGTGGACA 1000
Qy 1117 tctccatgcctgagtgacaactggtttgtccctgacactgcgcacagacagacagaatgac 1176
Db 1001 TCCCTCTGCTGGGATCAATAATATGGGCTCTCTGTGGTGGCCAGCCGACGAGAATGAT 1060
Qy 1177 ttcatccctcctgtctccacagtgacgggggcaggggtgtaggtgcatggttcaggatgag 1236
Db 1061 TTCAATCCCTCTGTGTTCAACGGTAACGTGGACCAAGGTCAATGGTCCACGATCAGAATGAG 1120
Qy 1237 cctgccttattgagatggttggttccaaacttgagggcctggcgtggagacacacacacacacac 1296
Db 1121 CCAGCCTTCATGATGATGGGGTTTCAATGTGCGTCCAGGTATGAGACCTCCATCAGC 1180
Qy 1297 atgagaagaagccctggacagcctcggaggaataatcacggcagctgactgagaatggt 1356
Db 1181 ATGAGGAAGGAGATGATGAGCGCTTCTGGGGCAGTTACAGCGACTGTACAGAGGATGGC 1240
Qy 1357 agcagtgcccggtcaagaaccttacccttccaaagtatacacacagcaggtgtgcatcac 1416
Db 1241 AGCGATGTGCAGTGCAAAATCTGTACTCATCTCGCTACACTGAACAGGCTGCATTCGT 1300
Qy 1417 tccctgtccaggagaacatgatcaagaagtgtggtgctgctacacatcttccacctaag 1476
Db 1301 TCCTGCTTTACGCTCAACATGGTAAGCGGTGCTCCTGTGCTCTTACTTTTATCCCTTA 1360

Qy 1477 cccaagggagttgagttctgtgaactaccgaaagcagagcctcctcgggctattgtctattat 1536
Db 1361 CCTGATGGAGCAGAGTATTGTGACTACACAAGACAGCTAGTTGGGGTACTGCTATTATAC 1420
Qy 1537 aaactgcagggcgcccttctccttgagacagcctgggctgtttctcocaagtgtcggaagcct 1596
Db 1421 AAACCTCCTTGTCTGAATTCAAAAGCTGATGTGCTGGGCTGCTTCCACAAATGTGGAAAGCT 1480
Qy 1597 tgtagtgtgatacaactacaaaactctcctcggcgtactcagcgttgcccaactctgtgaagtc 1656
Db 1481 TGCAAAATGACAGAATACCAGCTCTCAGCTGGATACTCTCGCTGGCCTTCTGCTGTCTCA 1540
Qy 1657 caggattggtatcttcgagatgctgctccttcagacaataatacaactattacaacaaaaga 1716
Db 1541 GAGGATTGGGTTTTTTTACATGCTTTTCAACAACAGACAATAATACATATACATCCAGAGG 1600
Qy 1717 aacggagttgcaagctcaacatcttctcagagagctggaactataaaactaattcggag 1776
Db 1601 AATGAGTTGCAAAAGTGAATATCTTTTTTGGAGAAATGGAACACAGACCAATGGGAG 1660
Qy 1777 tctccttctgcagatggttcagcctcctgtccaaacctggcagccagtcagtcgagcctgtgg 1836
Db 1661 TCACCAGCCTTTACAGTAGTACCCTCTTGTCTCAGCTTGGGAATCAGTGGAGTCTCTGG 1720
Qy 1837 ttggctcgtcgtgctctcctgtggtgagatggcggagctcatcttcgacctcctgctc 1896
Db 1721 TTTGGATCCTCTGTCTTCAATGATGAGGCTTGTGAGCTGATTTCTGGATTTCACTGTC 1780
Qy 1897 atcacacttctcatgctgctacgcgggttcgggagccgggtactgg 1941
Db 1781 ATCACTTTTCATCCTGGCCTTCCGCTGGTTCCGCTCCAAGCAGTGG 1825

Search completed: September 10, 2000, 23:02:51
Job time: 5759 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 22:15:07 ; Search time 97.02 Seconds
(without alignments)
8038.020 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtcaccacgcgtcg.....aaaaaaaaaaaaaaaa 3117

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0
Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	45	1.4	45	1 T99077	Probe Ena-A7 for a
C 2	45	1.4	45	1 T99078	Probe Ena-A9 for a
C 3	38	1.2	75	1 T99072	Alpha epithelial s
C 4	38	1.2	960	1 Q03293	Recombinant DNA en
C 5	38	1.2	1276	1 V29126	HPTP-1 coding sequ
C 6	38	1.2	3557	1 N60801	Human pro-growth h
C 7	38	1.2	3621	1 N60847	Human pre-prolacti
C 8	37	1.2	495	1 T74484	Pilosulin 2 precur
C 9	37	1.2	606	1 V73439	Flea saliva protei
C 10	37	1.2	606	1 V73440	Flea saliva protei
C 11	37	1.2	732	1 V59566	Human secreted pro
C 12	37	1.2	970	1 N50414	Sequence encoding
C 13	37	1.2	1001	1 T73851	Full length 3' UTR
C 14	37	1.2	1889	1 V69433	Human edg-6 cDNA #
C 15	37	1.2	2158	1 Q25156	Alpha-GalNac from
C 16	37	1.2	2702	1 V55872	Plant acetolactate
C 17	36	1.2	340	1 V51737	DNA encoding a hum
C 18	36	1.2	378	1 V90045	EST clone CW1550.
C 19	36	1.2	433	1 Q23343	Murine Cp-10. New
C 20	36	1.2	437	1 X22245	Human secreted pro
C 21	36	1.2	505	1 X04408	Human secreted pro
C 22	36	1.2	558	1 T64563	Mouse thymus and a
C 23	36	1.2	581	1 T63172	Trypanosoma cruzi
C 24	36	1.2	582	1 T64562	Human thymus and a
C 25	36	1.2	602	1 X20419	Human secreted pro
C 26	36	1.2	684	1 Q72956	Clone pSP65-Xa1-11
C 27	36	1.2	687	1 X30351	DNA encoding a hum
C 28	36	1.2	742	1 Q22005	Sequence encoding
C 29	36	1.2	780	1 T59912	Human transcriptio
C 30	36	1.2	794	1 N92416	Sequence of the S3
C 31	36	1.2	808	1 X04314	Human secreted pro
C 32	36	1.2	886	1 X51754	DNA encoding a hum
C 33	36	1.2	893	1 Q92529	P. communis (pear)

Human adult retina	1	V99723
Human secreted pro	1	X27330
DNA encoding a hum	1	X51728
RP-8 Programmed ce	1	O73760
Coding sequence fo	1	T35704
DNA encoding a hum	1	X51748
Human secreted pro	1	X34470
Human chemokine 2S	1	X45444
Human secreted pro	1	X30150
Plasmid pMG3C9 use	1	Q10377
Human secreted pro	1	X37464
Human secreted pro	1	X37460
DNA encoding a hum	1	X30366
P. tetragonolobus	1	T75765
Nucleotide sequenc	1	V59118
Human secreted pro	1	V43611
Human secreted pro	1	V62754
Clone W264 encodin	1	Q23028
Wheat leaf ADP-glu	1	O66589
Human sperm-specif	1	X06777
NF-YB. Expression	1	Q15361
Human secreted pro	1	V80740
Human transaldolas	1	V40647
Rabbit zona pelluc	1	X02709
Human transaldolas	1	X02709
Rabbit zpc cDNA. I	1	V64789
Original S. tubero	1	X27067
Gibberellin-20-ox	1	Q75359
Human secreted pro	1	X19486
Oil seed rape cyst	1	T90173
Human sperm-specif	1	X06779
Human foetal brain	1	V73000
Clone W282 encodin	1	Q23029
Human mitogen-acti	1	X07066
T cell replacing f	1	T14925
Plasmid pSP6K-mTRF	1	V64062
Human basigin I im	1	Q71341
Cystathionine gamm	1	T06480
Human TNF receptor	1	V34509
Human tumour necro	1	V83763
Gene encoding huma	1	V08856
Human secreted pro	1	V84616
Tyrosine Kinase re	1	Q27539
Oligodendrocyte-my	1	T30394
Xr1 coding sequenc	1	Q39085
Human secreted pro	1	X27383
Mouse neuronal PAS	1	V41257
Sequence of part o	1	Q43853
Macadamia integrif	1	V42316
Pea plastidial pho	1	V00731
A. thaliana sterol	1	V45981
Human c-myc far up	1	Q68910
Human c-myc far up	1	Q68910
p.mysA cDNA clone.	1	Q46672
cDNA from clone cr	1	V63189
Coding sequence fo	1	V44889
Human secreted pro	1	V59691
Human adult brain	1	V73003
Sequence encoding	1	Q83532
Homo sapiens secre	1	V40547
Cell membrane prot	1	V02308
Human N-arginine d	1	X21560
Sequence encoding	1	Q05326
Plasmid pfgH1 enco	1	Q03517
Prostate-tumour de	1	X27277
Sequence of new pl	1	N81634
Yellowtail cuna DN	1	T79857
Yellowtail/fin tu	1	T90019
Human secreted pro	1	V84428
Coding sequence of	1	V40888
Intron 8 of human	1	T96838
Human ADNF-III cDN	1	V49807

107	36	1..2	4978	1	Q74281	PLA2-I receptor en	180	35	1..1	1036	1	V84486	Human secreted pro
108	36	1..2	4978	1	T02986	Bovine phospholipase	181	35	1..1	1040	1	V02818	Human RHAMM genom
109	36	1..2	5668	1	X27262	Prostate-tumour de	182	35	1..1	1066	1	Q49943	Human anti-HBs lig
110	36	1..2	6254	1	V40887	Coding sequence of	183	35	1..1	1069	1	V84474	Human secreted pro
111	36	1..2	237326	1	V57903	Hereditary haemoch	184	35	1..1	1071	1	X20414	Human secreted pro
112	35	1..1	57	1	V27940	Human GAI3 primer	185	35	1..1	1097	1	V44860	Clone BG142_1 codi
113	35	1..1	70	1	T17035	Human mitochondria	186	35	1..1	1128	1	T86027	Meripilus giganteu
114	35	1..1	75	1	T17036	Human mitochondria	187	35	1..1	1180	1	V84347	Human Apo-2DCR cDN
115	35	1..1	85	1	V32418	Homo sapiens clone	188	35	1..1	1181	1	V59803	Human secreted pro
116	35	1..1	90	1	V37789	Analytical solid p	189	35	1..1	1196	1	V20812	Human haematopoiet
117	35	1..1	106	1	V00430	3' fragment of clo	190	35	1..1	1212	1	V59886	Human secreted pro
118	35	1..1	106	1	V11611	Homo sapiens adult	191	35	1..1	1215	1	X00620	Human secreted pro
119	35	1..1	160	1	T61985	Detection probe fo	192	35	1..1	1225	1	T33325	American elm chiti
120	35	1..1	179	1	V11620	Homo sapiens adult	193	35	1..1	1248	1	V58644	Prostate tumour sp
121	35	1..1	204	1	T61988	Detection probe fo	194	35	1..1	1248	1	V61249	cDNA sequence of p
122	35	1..1	207	1	Q50399	GPf - 1E, part of	195	35	1..1	1261	1	X04382	Human secreted pro
123	35	1..1	259	1	V00423	3' fragment of clo	196	35	1..1	1276	1	T42071	Human nuclear prol
124	35	1..1	263	1	T85891	Malassezia fungus	197	35	1..1	1285	1	Q86744	TRMP-3 metalloprot
125	35	1..1	283	1	V89992	EST clone CW762, N	198	35	1..1	1302	1	Q11112	MHC class II antig
126	35	1..1	308	1	V02985	Human BF245_1 3' p	199	35	1..1	1305	1	V26270	Trichomonas vagina
127	35	1..1	317	1	V11421	Human secreted pro	200	35	1..1	1305	1	V26272	Trichomonas vagina
128	35	1..1	308	1	V11421	Human secreted pro	201	35	1..1	1307	1	V63176	cDNA from clone dx
129	35	1..1	317	1	Q55664	AG1 toxin, New pep	202	35	1..1	1322	1	X04402	Human secreted pro
130	35	1..1	470	1	V58579	Prostate tumour sp	203	35	1..1	1328	1	V58754	Human secreted pro
131	35	1..1	470	1	V61312	cDNA sequence of p	204	35	1..1	1337	1	X04345	Human secreted pro
132	35	1..1	509	1	V58672	Prostate tumour sp	205	35	1..1	1355	1	N80457	Sequence of cDNA e
133	35	1..1	509	1	V61342	Extended cDNA sequ	206	35	1..1	1361	1	X51714	DNA encoding a hum
134	35	1..1	553	1	V31993	BS106 polynucleoti	207	35	1..1	1380	1	T51051	Human anine recept
135	35	1..1	571	1	V04715	Human secreted pro	208	35	1..1	1386	1	V11855	Homo sapiens Tub I
136	35	1..1	572	1	V44301	Human secreted pro	209	35	1..1	1391	1	T07329	Human calcium sign
137	35	1..1	579	1	Q14800	Human chorionic go	210	35	1..1	1423	1	N90249	Cucumber mosaic vi
138	35	1..1	582	1	X37502	Human secreted pro	211	35	1..1	1426	1	T33115	Human placental ri
139	35	1..1	588	1	V84464	Human secreted pro	212	35	1..1	1426	1	T72272	Cucumber mosaic vi
140	35	1..1	600	1	X07427	Homo sapiens secre	213	35	1..1	1448	1	X37462	Human secreted pro
141	35	1..1	603	1	T91169	DNA encoding a hum	214	35	1..1	1467	1	V82457	Triticum sp. cyste
142	35	1..1	630	1	T45865	cDNA encoding C. h	215	35	1..1	1493	1	Q28411	Glia activating fa
143	35	1..1	631	1	V82801	ATG-1120 (allograf	216	35	1..1	1503	1	V29523	Homo sapiens PSP1
144	35	1..1	635	1	Q75445	Murine glycosylati	217	35	1..1	1512	1	T72172	Alzheimer's diseas
145	35	1..1	635	1	T38390	Murine glycosylati	218	35	1..1	1522	1	V46309	Human secreted pro
146	35	1..1	643	1	V84539	Human secreted pro	219	35	1..1	1525	1	V22704	Mouse recombinase
147	35	1..1	669	1	N90103	Human prealbumin c	220	35	1..1	1525	1	X00679	Human secreted pro
148	35	1..1	685	1	X27366	Human secreted pro	221	35	1..1	1527	1	Q35200	DHP-1 polypeptide
149	35	1..1	685	1	X37398	Human secreted pro	222	35	1..1	1534	1	T15995	Human interleukin-
150	35	1..1	687	1	X30410	DNA encoding a hum	223	35	1..1	1538	1	V84593	Human secreted pro
151	35	1..1	688	1	T16856	Integrin subunit b	224	35	1..1	1540	1	V82787	Clone gq38_1 isola
152	35	1..1	712	1	V59811	Human secreted pro	225	35	1..1	1547	1	X33448	Oryza sativa L. pi
153	35	1..1	713	1	Q48773	Aphrodite. DNA a	226	35	1..1	1554	1	X16675	Xenopus WA545 prot
154	35	1..1	713	1	Q49210	Hamster Aphrodite	227	35	1..1	1573	1	V43607	Human secreted pro
155	35	1..1	745	1	V82783	Clone dn809_5 isol	228	35	1..1	1576	1	Q49944	Human anti-HBs hea
156	35	1..1	755	1	Q03970	Clone P25b New pro	229	35	1..1	1607	1	T85878	Malassezia fungus
157	35	1..1	767	1	T15821	TBE virus strain N	230	35	1..1	1613	1	X26196	DNA sequence of gl
158	35	1..1	767	1	T44470	Tick-borne enceph	231	35	1..1	1621	1	V58584	Prostate tumour sp
159	35	1..1	801	1	N91234	cDNA clone 8F3 of	232	35	1..1	1621	1	V61199	Full length cDNA s
160	35	1..1	831	1	V84475	Human secreted pro	233	35	1..1	1637	1	V84578	Human secreted pro
161	35	1..1	832	1	X03363	DNA encoding a hum	234	35	1..1	1642	1	V16096	Tyrosine kinase as
162	35	1..1	857	1	Q46850	Recombinant human	235	35	1..1	1642	1	V40507	Human secreted pro
163	35	1..1	857	1	T14564	Human milk kappa-c	236	35	1..1	1651	1	Q80218	Human NDF-alpha2b
164	35	1..1	863	1	T91302	Human HI075-1 secr	237	35	1..1	1670	1	V46310	Human secreted pro
165	35	1..1	877	1	V04635	Human cytoplasmic	238	35	1..1	1686	1	T91311	Human J143-1 secre
166	35	1..1	885	1	N90091	Gene 519 cDNA New	239	35	1..1	1693	1	V69206	Canine ICAM-R poly
167	35	1..1	907	1	V59695	Human secreted pro	240	35	1..1	1720	1	X27439	Human secreted pro
168	35	1..1	940	1	T58281	Arabidopsis STZ po	241	35	1..1	1733	1	V81394	Human tumour antig
169	35	1..1	941	1	V84586	Human secreted pro	242	35	1..1	1755	1	Q71873	Sequence coding hu
170	35	1..1	954	1	X37477	Human secreted pro	243	35	1..1	1767	1	X51749	DNA encoding a hum
171	35	1..1	971	1	X37469	Human secreted pro	244	35	1..1	1781	1	Q46991	cDNA clone encodin
172	35	1..1	978	1	X51732	DNA encoding a hum	245	35	1..1	1781	1	V19328	Human ICAM-4 cDNA
173	35	1..1	980	1	X37466	Human secreted pro	246	35	1..1	1781	1	V11657	Human ICAM-R cDNA
174	35	1..1	981	1	V60576	ERAB protein codin	247	35	1..1	1781	1	V56348	Human ICAM-R cDNA
175	35	1..1	986	1	X51735	DNA encoding a hum	248	35	1..1	1781	1	V54926	cDNA encoding huma
176	35	1..1	996	1	Q89606	Human membrane pre	249	35	1..1	1781	1	V69125	Intercellular adhe
177	35	1..1	1013	1	Q12431	Fragment D of urat	250	35	1..1	1781	1	X21855	Human ICAM-R codin
178	35	1..1	1023	1	T70132	Max-interacting pr	251	35	1..1	1813	1	V70895	cDNA encoding an a
179	35	1..1	1024	1	V84444	Human secreted pro	252	35	1..1	1921	1	T08610	Yeast heterogene e

253	35	1.1	1965	1	Q11815	HT7 membrane prote	326	1.1	5769	1	V15004	Receptor protein t
254	35	1.1	1971	1	X10156	Sequence encoding	327	1.1	5894	1	T85474	hABC3 cDNA sequenc
255	35	1.1	1984	1	Q52252	Protein PRO271 cDN	328	1.1	5894	1	V16307	cDNA encoding huma
256	35	1.1	2001	1	T59700	PTH-like peptide D	329	1.1	5975	1	N91778	Rabbit skeletal mus
257	35	1.1	2025	1	V59809	Human secreted pro	330	1.1	5975	1	Q87978	Rabbit skeletal ca
258	35	1.1	2045	1	V34297	Human secreted pro	331	1.1	5975	1	T70228	Rabbit calcium cha
259	35	1.1	2065	1	T58164	Rice ADP-glucose p	332	1.1	5975	1	T96811	Rabbit skeletal ca
260	35	1.1	2065	1	T29062	BRCA1 modulator pr	333	1.1	6476	1	X24302	Murine mWRN gene (
261	35	1.1	2094	1	X19489	Human secreted cDN	334	1.1	6525	1	V16345	cDNA encoding full
262	35	1.1	2116	1	Q35987	Tomato hsp80 cDNA	335	1.1	6644	1	X33181	Base sequence of t
263	35	1.1	2132	1	X33816	Coding sequence fo	336	1.1	7372	1	X33182	Base sequence of t
264	35	1.1	2144	1	T29538	Homo sapiens PSPI-	337	1.1	7797	1	X33180	Cowpox virus bsr f
265	35	1.1	2152	1	T22002	Human MED1 endonuc	338	1.1	7859	1	Q12154	Shuttle vector pSW
266	35	1.1	2158	1	T15228	Tumor necrosis fac	339	1.1	7859	1	Q42465	PSW6 for expressio
267	35	1.1	2158	1	T94631	TNF-R1-DD ligand p	340	1.1	7984	1	Q25185	PSW6 expression ve
268	35	1.1	2171	1	V73006	Human foetal brain	341	1.1	7996	1	X33184	Base sequence of t
269	35	1.1	2177	1	N60063	Partial Factor VII	342	1.1	9080	1	V19580	Human factor III e
270	35	1.1	2189	1	N80944	Alpha-2 plasmin in	343	1.1	10660	1	Q84793	Spinocerebellar at
271	35	1.1	2209	1	T87427	Clone J5. Polynucl	344	1.1	11141	1	T15820	TBE virus strain N
272	35	1.1	2211	1	Q12190	Human sperm-specif	345	1.1	11141	1	T44469	Tick-borne encephal
273	35	1.1	2238	1	V24017	Human HBMBU14 codi	346	1.1	11517	1	Q26021	PSP6-SFV4 RNA tran
274	35	1.1	2243	1	Q13402	Human alpha-2 plas	347	1.1	12827	1	V09036	Equine arteritis v
275	35	1.1	2249	1	N92325	Human alpha-2-plas	348	1.1	15528	1	V09039	Equine arteritis v
276	35	1.1	2253	1	Q05998	Sequence encoding	349	1.1	29392	1	V15422	Mouse poly Ig rece
277	35	1.1	2256	1	Q99002	Human acetylcholin	350	1.1	40	1	Q25031	Oligonucleotide sp
278	35	1.1	2280	1	V52604	Human metalloprote	351	1.1	44	1	V09273	Nucleotide sequenc
279	35	1.1	2335	1	Q80216	Human proNDP-alpha	352	1.1	48	1	T04081	Trypsin inhibitory
280	35	1.1	2346	1	V59693	Human secreted pro	353	1.1	49	1	V02146	Human secreted pro
281	35	1.1	2424	1	N91177	Sequence encoding	354	1.1	49	1	T88080	3' portion of cDNA
282	35	1.1	2432	1	X00650	Human secreted pro	355	1.1	50	1	Q66922	Poly-dA 50mer prob
283	35	1.1	2435	1	X00707	Human secreted pro	356	1.1	50	1	V37933	Primer of the spec
284	35	1.1	2438	1	N60065	Factor IX/Factor V	357	1.1	56	1	Q52732	Sequence of oligo
285	35	1.1	2445	1	N60111	Sequence encoding	358	1.1	56	1	O52734	Sequence of oligo
286	35	1.1	2445	1	Q48393	Full-length foetal	359	1.1	61	1	Q78635	Human chondromodul
287	35	1.1	2541	1	X04343	Human secreted pro	360	1.1	61	1	V07498	Lelystad virus pri
288	35	1.1	2594	1	T02367	Melon aconitase cD	361	1.1	69	1	V02148	Human secreted pro
289	35	1.1	2605	1	V04699	Homo sapiens 20q13	362	1.1	69	1	T88081	3' portion of cDNA
290	35	1.1	2628	1	V01060	Human PKR gene. Sc	363	1.1	69	1	V21236	Homo sapiens clone
291	35	1.1	2634	1	X27062	S. tuberosum isoam	364	1.1	69	1	V21226	Homo sapiens clone
292	35	1.1	2643	1	T33855	Human urinary blad	365	1.1	70	1	V30923	Human secreted pro
293	35	1.1	2689	1	V49886	Human IgA nephropa	366	1.1	74	1	V99725	Human adult testis
294	35	1.1	2700	1	T97447	Pacific yew taxadi	367	1.1	77	1	V77194	Staphylococcus aur
295	35	1.1	2810	1	V17099	Human brain derive	368	1.1	79	1	V26706	Human novel secret
296	35	1.1	2880	1	X18193	hNBC3a coding sequ	369	1.1	79	1	V32414	Homo sapiens clone
297	35	1.1	3024	1	Q04049	Sequence of clone	370	1.1	80	1	V37197	Oligonucleotide se
298	35	1.1	3024	1	Q35260	ICAM-1 cDNA. Treat	371	1.1	80	1	V30930	Human secreted pro
299	35	1.1	3024	1	Q57397	Human ICAM-1. Inte	372	1.1	84	1	V04275	Secreted protein C
300	35	1.1	3024	1	T11672	Intracellular adhe	373	1.1	84	1	V09269	Nucleotide sequenc
301	35	1.1	3024	1	T47957	Human ICAM-1 (HL-6	374	1.1	85	1	T39470	Growth regulatory
302	35	1.1	3024	1	T62835	Human ICAM-1 cDNA.	375	1.1	85	1	V05720	Nucleotide sequenc
303	35	1.1	3024	1	V70950	Intracellular adhe	376	1.1	85	1	V37488	Human growth regul
304	35	1.1	3113	1	N80995	HL-cDNA clone enco	377	1.1	86	1	Q11760	Self-complementary
305	35	1.1	3114	1	V02800	Human RHAMM cDNA.	378	1.1	86	1	V11441	Human secreted pro
306	35	1.1	3176	1	T62124	Arabidopsis thalia	379	1.1	87	1	V05740	Nucleotide sequenc
307	35	1.1	3320	1	V56991	Maize starch phosph	380	1.1	90	1	Q11762	Self-complementary
308	35	1.1	3350	1	T18200	Infectious bursal	381	1.1	90	1	T47082	Synthetic DNA prob
309	35	1.1	3471	1	Q46673	P. mySB cDNA clone.	382	1.1	90	1	T47081	Synthetic DNA prob
310	35	1.1	3547	1	N60846	Plasmid sequence e	383	1.1	90	1	V31773	Synthetic oligonuc
311	35	1.1	3572	1	V67162	Drosophila kinesin	384	1.1	90	1	V54077	Nucleotide sequenc
312	35	1.1	3592	1	Q53529	Human Mannose-Bind	385	1.1	92	1	V11595	Homo sapiens adult
313	35	1.1	3692	1	T72107	Rat semaphorin 2 g	386	1.1	92	1	V61480	Human secreted pro
314	35	1.1	3798	1	Q42424	ADA2 DNA. Transcri	387	1.1	94	1	V05728	Nucleotide sequenc
315	35	1.1	3912	1	X18192	hNBC3 coding seque	388	1.1	98	1	Q11764	Self-complementary
316	35	1.1	3933	1	T03885	Human mucosal lymph	389	1.1	98	1	X00174	Porcine reproducti
317	35	1.1	4180	1	N91773	Rat androgen recep	390	1.1	99	1	T91300	Human M97-2 secret
318	35	1.1	4364	1	Q63730	Protein tyrosine k	391	1.1	100	1	V68827	DNA molecule encod
319	35	1.1	4589	1	X06068	p150.95 alpha subu	392	1.1	101	1	V00420	3' fragment of clo
320	35	1.1	4765	1	X09003	Mouse semaphorin r	393	1.1	112	1	V41441	Nucleotide sequenc
321	35	1.1	4780	1	X09004	Mouse semaphorin r	394	1.1	114	1	T39472	Growth regulatory
322	35	1.1	4985	1	X00463	Human type IX aden	395	1.1	114	1	V37490	Human growth regul
323	35	1.1	5555	1	Q64898	HER4 with alternat	396	1.1	117	1	V21223	Homo sapiens clone
324	35	1.1	5555	1	T18533	Receptor tyrosine	397	1.1	119	1	V44858	Clone Ad1.1 3' seq
325	35	1.1	5613	1	Q63241	Crucifer 1-aminoc	398	1.1	120	1	Q52429	Human subtelomeric

399	34	1.1	120	1	Q87011	Subtelomeric cDNA	472	34	1.1	570	1	T70048	Cotton fibre speci
400	34	1.1	127	1	T96338	Subtelomeric cDNA	473	34	1.1	572	1	X07182	Soybean threonine
401	34	1.1	120	1	T11604	Homo sapiens adult	474	34	1.1	577	1	V59806	Human secreted pro
402	34	1.1	140	1	T76781	Staphylococcus aur	475	34	1.1	579	1	V82799	ATG-1100 (allograf
403	34	1.1	141	1	T06012	Immunodominant fra	476	34	1.1	585	1	X37386	Human secreted pro
404	34	1.1	141	1	T34595	Probe for detectin	477	34	1.1	596	1	Q10488	Fusion protein enc
405	34	1.1	144	1	V21239	Homo sapiens clone	478	34	1.1	596	1	V00462	Human interleukin-
406	34	1.1	191	1	T13475	Capture probe for	479	34	1.1	600	1	Q72739	T. ovis vaccine ca
407	34	1.1	198	1	Q42784	Ligand-induced gen	480	34	1.1	600	1	T17934	Taenia ovis antige
408	34	1.1	201	1	V00418	3' fragment of clo	481	34	1.1	604	1	X37373	Human secreted pro
409	34	1.1	202	1	V41451	Nucleotide sequenc	482	34	1.1	609	1	T62617	Cotton fibre speci
410	34	1.1	206	1	V09116	3' nucleotide sequ	483	34	1.1	615	1	V02881	Human HMGI-C aberr
411	34	1.1	218	1	V09114	3' nucleotide sequ	484	34	1.1	618	1	V58362	Coding sequence fo
412	34	1.1	222	1	Q46071	Sequence downstrea	485	34	1.1	625	1	V40885	Coding sequence of
413	34	1.1	227	1	T89601	Hepatitis C virus	486	34	1.1	625	1	X19927	Lily LGC1 nucleoti
414	34	1.1	240	1	T76782	Hepatitis C virus	487	34	1.1	633	1	V09282	Nucleotide sequenc
415	34	1.1	249	1	T89608	Hepatitis C virus	488	34	1.1	634	1	X30403	DNA encoding a hum
416	34	1.1	257	1	T89599	Hepatitis C virus	489	34	1.1	644	1	X00684	Human secreted pro
417	34	1.1	260	1	T89601	Hepatitis C virus	490	34	1.1	645	1	X04348	Human secreted pro
418	34	1.1	270	1	T89602	Hepatitis C virus	491	34	1.1	660	1	X37406	Human secreted pro
419	34	1.1	286	1	X30353	DNA encoding a hum	492	34	1.1	667	1	Q61566	Human SOD. Recombi
420	34	1.1	289	1	Q53450	Sequence of the 3'	493	34	1.1	667	1	Q61573	Human SOD. Recombi
421	34	1.1	292	1	V21153	3' nucleotide port	494	34	1.1	667	1	Q61575	Hybrid human/plig S
422	34	1.1	296	1	X25862	Clone PEGen-44 seq	495	34	1.1	667	1	X04334	Human secreted pro
423	34	1.1	314	1	V09108	3' nucleotide sequ	496	34	1.1	688	1	Q61572	Human SOD. Recombi
424	34	1.1	321	1	T89598	Hepatitis C virus	497	34	1.1	688	1	V34216	Human secreted pro
425	34	1.1	334	1	X25857	Clone PEGen-24 seq	498	34	1.1	699	1	O58663	Human superoxide d
426	34	1.1	337	1	X25860	Clone PEGen-42 seq	499	34	1.1	700	1	Q22352	Pl4(?) allergen cl
427	34	1.1	340	1	Q99393	Rat allograft infl	500	34	1.1	700	1	Q89191	Birch pollen Pl4 a
428	34	1.1	340	1	V89136	EST clone BR309. N	501	34	1.1	700	1	V34301	Human secreted pro
429	34	1.1	342	1	Q59382	Human brain Expres	502	34	1.1	704	1	V35167	Thuga plicata diri
430	34	1.1	345	1	V26683	Human novel secret	503	34	1.1	704	1	V84441	Human secreted pro
431	34	1.1	347	1	Q27482	T23 CDNA. Stamen-s	504	34	1.1	719	1	V40501	Human secreted pro
432	34	1.1	350	1	Q98380	Rat allograft infl	505	34	1.1	722	1	X06786	Human adult uterus
433	34	1.1	356	1	T89597	Hepatitis C virus	506	34	1.1	724	1	V58515	3' fragment of pro
434	34	1.1	372	1	N80489	Cowpea trypsin inh	507	34	1.1	724	1	V61281	3' cDNA sequence o
435	34	1.1	373	1	X30409	DNA encoding a hum	508	34	1.1	728	1	T85876	Malassezia fungus
436	34	1.1	375	1	T84939	Human prostate pro	509	34	1.1	730	1	V00421	Clone B219.2. Isol
437	34	1.1	382	1	Q60162	Human brain Expres	510	34	1.1	737	1	V34270	Human secreted pro
438	34	1.1	399	1	T13044	Cotton fibre-speci	511	34	1.1	740	1	V23915	Plant CCR enzyme D
439	34	1.1	399	1	T62620	Cotton fibre cell-	512	34	1.1	741	1	V23878	Plant CCR enzyme D
440	34	1.1	399	1	T62620	Cotton fibre speci	513	34	1.1	742	1	V00437	Clone H905_107. Is
441	34	1.1	399	1	T70051	Cotton fibre speci	514	34	1.1	748	1	Q10376	Plasmid pMG4B12 us
442	34	1.1	413	1	X27371	Human secreted pro	515	34	1.1	755	1	X37467	Human secreted pro
443	34	1.1	423	1	X25073	Tobacco NADP-ferro	516	34	1.1	756	1	Q73729	Rape abscission/de
444	34	1.1	427	1	T85904	Malassezia fungus	517	34	1.1	760	1	V84625	Human secreted pro
445	34	1.1	429	1	V09097	Nucleotide sequenc	518	34	1.1	766	1	N40162	Sequence of prepro
446	34	1.1	429	1	V15712	Clone BV239_3 nucl	519	34	1.1	769	1	T79095	Full length human
447	34	1.1	443	1	N70773	Sequence encoding	520	34	1.1	772	1	V58363	Coding sequence fo
448	34	1.1	458	1	Q90525	Rat SIII 15 kDa su	521	34	1.1	773	1	X37388	Human secreted pro
449	34	1.1	458	1	T59893	Rat RNA polymerase	522	34	1.1	775	1	X26148	Human pan-s/tk-1A
450	34	1.1	464	1	V38682	Mus musculus SOCS1	523	34	1.1	776	1	X30318	DNA encoding a hum
451	34	1.1	467	1	V64423	Mouse developing l	524	34	1.1	778	1	X30336	DNA encoding a hum
452	34	1.1	472	1	V84530	Human secreted pro	525	34	1.1	779	1	Q79737	Flower style-speci
453	34	1.1	479	1	Q32422	Homo sapiens clone	526	34	1.1	779	1	T64554	Tomato S-ribonucle
454	34	1.1	487	1	Q11686	Sequence encoding	527	34	1.1	785	1	V84473	Human secreted pro
455	34	1.1	487	1	X30156	Human secreted pro	528	34	1.1	787	1	T29008	Parietaria allerg
456	34	1.1	492	1	X22223	Human secreted pro	529	34	1.1	789	1	V58509	3' fragment of pro
457	34	1.1	513	1	X22214	Human secreted pro	530	34	1.1	789	1	V61275	3' cDNA sequence o
458	34	1.1	515	1	T13826	CHL1 neural cell a	531	34	1.1	791	1	X19493	Human secreted pro
459	34	1.1	524	1	V73005	Human adult testis	532	34	1.1	796	1	Q27948	CA455. Anther-spec
460	34	1.1	525	1	V41453	Nucleotide sequenc	533	34	1.1	801	1	V15416	Human chemokine MI
461	34	1.1	528	1	V02891	Human HMGI-C aberr	534	34	1.1	804	1	V59625	Human secreted pro
462	34	1.1	536	1	Q53448	Sequence encoding	535	34	1.1	809	1	V59739	Human secreted pro
463	34	1.1	546	1	X37513	Human secreted pro	536	34	1.1	812	1	T85877	Malassezia fungus
464	34	1.1	550	1	V41916	Nucleotide sequenc	537	34	1.1	819	1	V63198	cDNA from clone ga
465	34	1.1	553	1	X51720	DNA encoding a hum	538	34	1.1	824	1	X37387	Human secreted pro
466	34	1.1	556	1	V66764	Pathogen response	539	34	1.1	826	1	Q79736	Flower style-speci
467	34	1.1	562	1	V64661	Seq ID 29 from DEI	540	34	1.1	826	1	T64553	Tomato S-ribonucle
468	34	1.1	568	1	V53256	ETH precursor anti	541	34	1.1	836	1	X27402	Human secreted pro
469	34	1.1	568	1	V34054	Antisense full len	542	34	1.1	836	1	X04760	DNA encoding a coc
470	34	1.1	570	1	T13041	Cotton fibre-speci	543	34	1.1	859	1	T62944	Human eosinocyte C
471	34	1.1	570	1	T30258	Cotton fibre cell-	544	34	1.1	860	1	Q22007	Sequence encoding

545	34	1.1	861	1	Q99552	Thrombopoietin cod	618	34	1.1	1149	1	X27384	Human secreted pro
546	34	1.1	861	1	V40622	DN722.2 cDNA clone	619	34	1.1	1153	1	V59541	Human secreted pro
547	34	1.1	865	1	V48123	Nucleotide sequence	620	34	1.1	1153	1	V55748	Human secreted pro
548	34	1.1	865	1	V84470	Human secreted pro	621	34	1.1	1154	1	T39048	cDNA encoding cell
549	34	1.1	865	1	X30358	DNA encoding a hum	622	34	1.1	1160	1	V40008	Human cell death a
550	34	1.1	873	1	V35164	Thuga plicata diri	623	34	1.1	1163	1	Q35198	ppcDHP71 coding se
551	34	1.1	879	1	Q71632	Apo-B RNA editing	624	34	1.1	1164	1	V07654	Nucleotide sequenc
552	34	1.1	880	1	T91714	Rat SmLIM/CRP2 cDN	625	34	1.1	1165	1	N60391	Sequence encoding
553	34	1.1	882	1	T72173	Alzheimer's diseas	626	34	1.1	1166	1	T60972	Mouse lipopolysacc
554	34	1.1	887	1	Q05160	Fragment of plasm	627	34	1.1	1169	1	V99983	3' untranslated re
555	34	1.1	887	1	Q06743	Insert from ptiGH6	628	34	1.1	1172	1	Q58689	Protein allergen o
556	34	1.1	889	1	V19304	Human p26 cDNA seq	629	34	1.1	1173	1	Q29636	Hepatitis C virus
557	34	1.1	890	1	X04376	Human secreted pro	630	34	1.1	1173	1	Q43894	NANB hepatitis vir
558	34	1.1	903	1	T47670	Human G protein ga	631	34	1.1	1173	1	Q43896	NANB hepatitis vir
559	34	1.1	906	1	Q80222	Human NDF-beta3 cl	632	34	1.1	1173	1	Q43897	NANB hepatitis vir
560	34	1.1	911	1	T27409	Human secreted pro	633	34	1.1	1174	1	T39050	cDNA encoding cell
561	34	1.1	916	1	T65651	cDNA encoding anti	634	34	1.1	1174	1	V39096	Monocomponent endo
562	34	1.1	916	1	V62000	R. prolixus NO-r p	635	34	1.1	1181	1	V81616	Cucumber peroxidase
563	34	1.1	917	1	X27343	Human secreted pro	636	34	1.1	1181	1	V62867	Cucumber peroxidase
564	34	1.1	921	1	V40521	Homo sapiens CH27-	637	34	1.1	1189	1	V41452	Nucleotide sequenc
565	34	1.1	934	1	X04394	Human secreted pro	638	34	1.1	1210	1	V84425	Human secreted pro
566	34	1.1	936	1	T69171	Trypanosoma cruzi	639	34	1.1	1213	1	T61895	Human 14-3-3 beta
567	34	1.1	936	1	X51702	DNA encoding a hum	640	34	1.1	1215	1	T76885	Arabidopsis floral
568	34	1.1	938	1	V27142	Novel haemopoietin	641	34	1.1	1215	1	V58306	Arabidopsis thalia
569	34	1.1	941	1	N30062	Sequence of a modi	642	34	1.1	1218	1	T48447	Human TGF-beta-lik
570	34	1.1	943	1	V59535	Human secreted pro	643	34	1.1	1223	1	X25130	Soybean isoflavone
571	34	1.1	946	1	X37468	Human secreted pro	644	34	1.1	1227	1	T48446	Human TGF-beta-lik
572	34	1.1	953	1	T03910	Sequence encoding	645	34	1.1	1228	1	T48449	Human TGF-beta-lik
573	34	1.1	957	1	Q04441	Aequorin gene Bios	646	34	1.1	1229	1	Q38229	Sequence of rye gr
574	34	1.1	959	1	N81534	pA0440 aequorin ge	647	34	1.1	1229	1	Q85932	cDNA encoding Lol
575	34	1.1	961	1	Q13151	Heparin-binding gr	648	34	1.1	1230	1	X22276	Human secreted pro
576	34	1.1	966	1	T87430	Clone H174-10. Pol	649	34	1.1	1230	1	X07565	Homo sapiens fetal
577	34	1.1	966	1	X04360	Human secreted pro	650	34	1.1	1232	1	V40526	Homo sapiens C2268
578	34	1.1	975	1	V61499	Bak binding protei	651	34	1.1	1233	1	N91576	Sequence of clone
579	34	1.1	985	1	N90529	Polydeoxyribonucle	652	34	1.1	1239	1	Q23000	Sequence encoding
580	34	1.1	987	1	X04331	Human secreted pro	653	34	1.1	1239	1	Q43439	Lol pi cDNA clone
581	34	1.1	990	1	X37471	Human secreted pro	654	34	1.1	1240	1	X06787	Human adult trache
582	34	1.1	997	1	X52274	Protein PRO244 cDN	655	34	1.1	1242	1	T13809	Murine VWF186 cDNA
583	34	1.1	999	1	V80678	Human DNAX toll-li	656	34	1.1	1248	1	V11859	Homo sapiens Tub I
584	34	1.1	1001	1	X33229	Wheat viviparous l	657	34	1.1	1252	1	V32416	Homo sapiens clone
585	34	1.1	1034	1	Q12340	Caldesmon-like pol	658	34	1.1	1255	1	X37455	Human secreted pro
586	34	1.1	1015	1	X00713	Human secreted pro	659	34	1.1	1263	1	V99722	Human adult retina
587	34	1.1	1016	1	T29401	Lipase coding sequ	660	34	1.1	1265	1	V68898	DNA molecule encod
588	34	1.1	1020	1	X00681	Human secreted pro	661	34	1.1	1285	1	Q38028	CAD cDNA isolated
589	34	1.1	1025	1	X30319	DNA encoding a hum	662	34	1.1	1296	1	T04580	Human IFNAB-BPI. N
590	34	1.1	1032	1	X30155	Human secreted pro	663	34	1.1	1302	1	V99985	Nucleotide sequenc
591	34	1.1	1035	1	T60568	Cotton fibrous tis	664	34	1.1	1307	1	X30355	DNA encoding a hum
592	34	1.1	1041	1	Q54651	T84.12 light chain	665	34	1.1	1310	1	V40509	Human secreted pro
593	34	1.1	1041	1	Q54653	T84.12 L4-12-1 lig	666	34	1.1	1313	1	V59612	Human secreted pro
594	34	1.1	1045	1	Q79671	Sequence encoding	667	34	1.1	1315	1	T65022	Human nuclear prol
595	34	1.1	1047	1	Q30001	HCNP precursor gen	668	34	1.1	1315	1	V73311	Natural killer cel
596	34	1.1	1047	1	Q58685	Rat hippocampal ch	669	34	1.1	1316	1	T99543	Human chemokine re
597	34	1.1	1049	1	V43610	Human secreted pro	670	34	1.1	1322	1	V43616	Human secreted pro
598	34	1.1	1054	1	X00618	Human secreted pro	671	34	1.1	1323	1	T51738	Orange pectin meth
599	34	1.1	1065	1	X30407	DNA encoding a hum	672	34	1.1	1323	1	V64073	Orange pectin meth
600	34	1.1	1078	1	Q25785	Protein synthesis	673	34	1.1	1325	1	Q31873	Cyclin D1 gene. Re
601	34	1.1	1080	1	T04682	Human superoxide-d	674	34	1.1	1325	1	Q53197	Human cyclin D1. N
602	34	1.1	1083	1	V24126	Homo sapiens BE2 s	675	34	1.1	1325	1	V20473	Human cyclin D1 on
603	34	1.1	1090	1	X37521	Human secreted pro	676	34	1.1	1340	1	T85980	Human neuro-D gene
604	34	1.1	1096	1	X22275	Human secreted pro	677	34	1.1	1342	1	V62449	Human neurotensin
605	34	1.1	1113	1	Q29633	Hepatitis C virus	678	34	1.1	1348	1	Q78278	Human interleukin-
606	34	1.1	1114	1	V34315	Human secreted pro	679	34	1.1	1348	1	X24749	Human interleukin-
607	34	1.1	1118	1	Q20263	Deg-1 gene. DNA se	680	34	1.1	1364	1	T48099	Human interleukin-
608	34	1.1	1123	1	V34290	Human secreted pro	681	34	1.1	1364	1	V15228	Heterodimeric huma
609	34	1.1	1124	1	Q03669	Papaya ringspot vi	682	34	1.1	1364	1	V42536	Human interleukin-
610	34	1.1	1129	1	X27340	Human secreted pro	683	34	1.1	1367	1	X51727	DNA encoding a hum
611	34	1.1	1136	1	N70435	Sequence encoding	684	34	1.1	1370	1	X19886	Pancreas derived p
612	34	1.1	1139	1	V08813	Gene No. 3 encodin	685	34	1.1	1371	1	V17829	Homo sapiens pancr
613	34	1.1	1141	1	T13810	Murine VRF167 cDNA	686	34	1.1	1371	1	X19885	Pancreas derived p
614	34	1.1	1142	1	X51744	DNA encoding a hum	687	34	1.1	1373	1	V52969	Human G1 protein i
615	34	1.1	1143	1	V59619	Human secreted pro	688	34	1.1	1376	1	Q56747	Ryegrass Lol pv al
616	34	1.1	1146	1	X07000	Xenopus Frazzled p	689	34	1.1	1376	1	X04325	Human secreted pro
617	34	1.1	1147	1	V69620	Human secreted pro	690	34	1.1	1378	1	V59706	Human secreted pro

691	34	1.1	1384	1	X37442	Human secreted pro	764	34	1.1	1606	1	Q20264	Mec-4 gene. DNA se
692	34	1.1	1393	1	X38027	CAD cDNA isolated	765	34	1.1	1607	1	Q38934	mec-4 gene. Isolat
693	34	1.1	1393	1	Q66241	Alpha-galactosidas	766	34	1.1	1620	1	V43793	Rodent chemokine r
694	34	1.1	1393	1	Q77861	Human alpha-galact	767	34	1.1	1624	1	V17054	Glucosyl transfera
695	34	1.1	1394	1	Q05879	Placenta-specific	768	34	1.1	1627	1	V40486	Human secreted pro
696	34	1.1	1396	1	Q24371	Natural killer cel	769	34	1.1	1631	1	T72795	Beta-1 integrin mo
697	34	1.1	1400	1	Q06354	Recombinant urokin	770	34	1.1	1637	1	V69761	EDG-1-like G-prote
698	34	1.1	1400	1	T29191	Human urokinase pi	771	34	1.1	1640	1	Q50575	Asparaginylendopep
699	34	1.1	1400	1	X19823	Urokinase-type pla	772	34	1.1	1641	1	T15758	Human interleukin-
700	34	1.1	1404	1	V84622	Human secreted pro	773	34	1.1	1646	1	V69038	DNA molecule encod
701	34	1.1	1406	1	V61486	Human secreted pro	774	34	1.1	1653	1	X00699	Human secreted pro
702	34	1.1	1407	1	N60374	Sequence encoding	775	34	1.1	1654	1	T64943	cDNA encoding a no
703	34	1.1	1407	1	Q32515	Sequence encoding	776	34	1.1	1654	1	T93210	Human neurotactin
704	34	1.1	1407	1	Q71647	Human phospholipas	777	34	1.1	1657	1	T30360	Human p57 coding s
705	34	1.1	1407	1	T89675	Human phospholipas	778	34	1.1	1659	1	T31800	Maize adenylosuccl
706	34	1.1	1410	1	V15022	Hyoscynamine 6 beta	779	34	1.1	1668	1	X00664	Human secreted pro
707	34	1.1	1414	1	V33888	H. tuberosus CYCD3;	780	34	1.1	1670	1	V26557	Human IP-10/Mig re
708	34	1.1	1420	1	T72167	Alzheimer's diseas	781	34	1.1	1679	1	X00719	Human secreted pro
709	34	1.1	1420	1	X27321	Human secreted pro	782	34	1.1	1682	1	Q91763	Human Rad51 gene,
710	34	1.1	1425	1	T35220	Cytoplasmic antipr	783	34	1.1	1686	1	V73012	Human adult brain
711	34	1.1	1434	1	X37452	Human secreted pro	784	34	1.1	1687	1	V01072	Bovine p58 gene. S
712	34	1.1	1441	1	T48444	Human TGF-beta-lik	785	34	1.1	1694	1	V33190	Secreted protein B
713	34	1.1	1448	1	X00631	Human secreted pro	786	34	1.1	1694	1	X07431	Homo sapiens secre
714	34	1.1	1453	1	X00573	Asparaginylendopep	787	34	1.1	1699	1	V34294	Human secreted pro
715	34	1.1	1454	1	V23913	Plant CAD enzyme D	788	34	1.1	1700	1	V11858	Mus musculus Tub I
716	34	1.1	1461	1	T01471	Mammalian mast cel	789	34	1.1	1710	1	Q47966	Rape acyl-ACP thio
717	34	1.1	1469	1	V84120	Pentaclethra macro	790	34	1.1	1711	1	V43617	Human secreted pro
718	34	1.1	1474	1	T90174	Oil seed rape cyst	791	34	1.1	1713	1	T85993	Maize 5-enolpyruvy
719	34	1.1	1474	1	V23914	Plant CAD enzyme D	792	34	1.1	1713	1	T86000	Maize 5-enolpyruvy
720	34	1.1	1483	1	T27147	Chalcone synthase	793	34	1.1	1713	1	V19277	Maize 5-enolpyruvy
721	34	1.1	1487	1	X00698	Human secreted pro	794	34	1.1	1717	1	V08170	MMP19 coding seque
722	34	1.1	1490	1	Q75361	Gibberellin-20-ox	795	34	1.1	1718	1	X26151	3' CDNA sequence o
723	34	1.1	1493	1	Q04492	Sequence encoding	796	34	1.1	1721	1	T38129	Nuclear envelope-a
724	34	1.1	1493	1	X33994	Human osteopontin	797	34	1.1	1721	1	X07566	Homo sapiens fetal
725	34	1.1	1494	1	V41444	Nucleotide sequenc	798	34	1.1	1728	1	V38990	Human receptor for
726	34	1.1	1502	1	T68721	Human mitogen-acti	799	34	1.1	1730	1	T61340	Batten disease gen
727	34	1.1	1502	1	V10099	Human MAP kinase p	800	34	1.1	1731	1	T61348	Batten disease gen
728	34	1.1	1503	1	X32253	Protein PRO272 cDN	801	34	1.1	1731	1	T61333	Batten disease gen
729	34	1.1	1505	1	T72177	Alzheimer's diseas	802	34	1.1	1732	1	T61335	Batten disease gen
730	34	1.1	1505	1	V40522	Homo sapiens CI542	803	34	1.1	1732	1	T61336	Batten disease gen
731	34	1.1	1507	1	X00656	Human secreted pro	804	34	1.1	1732	1	T61337	Batten disease gen
732	34	1.1	1508	1	V07665	Maize Bx1 cDNA (DI	805	34	1.1	1732	1	T61338	Batten disease gen
733	34	1.1	1509	1	V59679	Human secreted pro	806	34	1.1	1732	1	T61341	Batten disease gen
734	34	1.1	1510	1	Q03740	Human myb related	807	34	1.1	1732	1	T61343	Batten disease gen
735	34	1.1	1511	1	T29032	Endoglucanase (S8	808	34	1.1	1732	1	T61344	Batten disease gen
736	34	1.1	1514	1	Q05981	Insert from pTPRLI	809	34	1.1	1732	1	T61345	Batten disease gen
737	34	1.1	1515	1	V84440	Human secreted pro	810	34	1.1	1732	1	T61346	Batten disease gen
738	34	1.1	1536	1	Q94111	mML genomic DNA. T	811	34	1.1	1732	1	T61347	Batten disease gen
739	34	1.1	1537	1	V59804	Human secreted pro	812	34	1.1	1732	1	T61332	Batten disease gen
740	34	1.1	1538	1	T63302	A. chrysoenum O-a	813	34	1.1	1732	1	T61306	Batten disease cDN
741	34	1.1	1541	1	Q35199	ppcDHP23 coding se	814	34	1.1	1733	1	T61339	Batten disease gen
742	34	1.1	1549	1	X24921	Guman ras carboxy-	815	34	1.1	1733	1	T61342	Batten disease gen
743	34	1.1	1553	1	T68695	Strawberry inverta	816	34	1.1	1734	1	T61334	Batten disease gen
744	34	1.1	1553	1	V61285	Human RCE1 (hRCE1)	817	34	1.1	1737	1	Q29506	New platelet facto
745	34	1.1	1558	1	Q10896	Encodes Xenopus Bo	818	34	1.1	1737	1	Q80521	Human monocyte PP4
746	34	1.1	1558	1	T28255	Survival motor neu	819	34	1.1	1737	1	Q90907	Chemokine superfam
747	34	1.1	1558	1	X22942	Human cathepsin Y	820	34	1.1	1738	1	Q56241	Sequence encoding
748	34	1.1	1559	1	T02714	MART-1 melanoma an	821	34	1.1	1741	1	T32830	Wheat adenylosuccl
749	34	1.1	1560	1	T18828	Human survival mot	822	34	1.1	1747	1	Q20260	ERK1. DNA encoding
750	34	1.1	1562	1	Q42769	Ligand-induced gen	823	34	1.1	1748	1	Q86327	Arabidopsis thalia
751	34	1.1	1564	1	T22231	Human secreted pro	824	34	1.1	1749	1	V21800	Human spleen-deriv
752	34	1.1	1569	1	N90111	Recombinant activa	825	34	1.1	1753	1	V59556	Human secreted pro
753	34	1.1	1575	1	N90112	Anticoagulant PP4	826	34	1.1	1758	1	T32325	soluble starch syn
754	34	1.1	1577	1	T90169	Oil seed rape cyst	827	34	1.1	1761	1	V41445	Nucleotide sequenc
755	34	1.1	1579	1	V64590	Tobacco chlI DNA.	828	34	1.1	1761	1	X20412	Human secreted pro
756	34	1.1	1582	1	T18831	Human survival mot	829	34	1.1	1769	1	T85973	Human NB Phox 3' u
757	34	1.1	1582	1	T28259	Survival motor neu	830	34	1.1	1770	1	X52242	Protein PRO233 cDN
758	34	1.1	1583	1	V58758	Human secreted pro	831	34	1.1	1772	1	V63174	cDNA from clone dd
759	34	1.1	1586	1	Q7967	Rape acyl-ACP thio	832	34	1.1	1776	1	T91744	Tobacco calcium/ca
760	34	1.1	1587	1	X06840	Arabidopsis branch	833	34	1.1	1780	1	V04269	Rat FRAG1 coding s
761	34	1.1	1594	1	V62755	Human secreted pro	834	34	1.1	1793	1	Q15139	Clone pACC6 encod
762	34	1.1	1604	1	V04252	Wild tomato acyltr	835	34	1.1	1801	1	N70524	Sequence of cDNA 1
763	34	1.1	1605	1	V59687	Human secreted pro	836	34	1.1	1813	1	Q88760	Human ubiquitous n

837	1.1	1817	1	V04064	Human RNA-binding	910	1.1	2120	1	X33812	Coding sequence fo
838	34	1821	1	V59590	Human secreted pro	911	34	2123	1	T59701	PTH-like peptide D
839	34	1824	1	Q47842	Flavonoid-3',5'-hy	912	34	2143	1	Q90242	Full length Sorghu
840	34	1829	1	V26298	Phosphate starvati	913	34	2153	1	V84565	Human secreted pro
841	34	1830	1	Q06612	Human secreted pro	914	34	2159	1	V57200	Human secreted pro
842	34	1834	1	Q04690	Encodes Mammalian	915	34	2159	1	Q67223	Mouse p55Nuc. Ant1
843	34	1835	1	V27245	Human CIAR1 coding	916	34	2196	1	V59663	Human secreted pro
844	34	1836	1	V33461	Human T1-receptor	917	34	2200	1	Q11010	Fibulin C. Purifie
845	34	1837	1	T33852	Rat vas deferens p	918	34	2205	1	V04680	Human presenilin 1
846	34	1841	1	Q91311	Human Fas-L gene.	919	34	2214	1	V40524	Homo sapiens COL22
847	34	1841	1	X21574	Human FasL protein	920	34	2218	1	V84503	Human secreted pro
848	34	1841	1	X24879	Human Fas ligand (921	34	2218	1	X02887	Rat mACHR-6 cDNA.
849	34	1841	1	V24018	Human BHF1 coding	922	34	2223	1	Q39051	Tomato acid invert
850	34	1848	1	V84588	Human secreted pro	923	34	2229	1	Q23462	Human B-raf protei
851	34	1849	1	T74042	Soybean thiol prot	924	34	2230	1	V15704	Tomato ACC synthas
852	34	1858	1	T72175	Alzheimer's diseas	925	34	2233	1	V10120	Human retinoid rec
853	34	1863	1	N60399	Sequence of the pr	926	34	2239	1	V01527	Wheat soluble star
854	34	1864	1	V59630	Human secreted pro	927	34	2255	1	N60228	Sequence of the br
855	34	1866	1	V84466	Human secreted pro	928	34	2255	1	Q84698	Human IFN-gamma ac
856	34	1868	1	T73502	Poplar 1-aminocycl	929	34	2264	1	X20496	Human secreted pro
857	34	1868	1	V06408	Epidermal growth f	930	34	2265	1	X20435	Human secreted pro
858	34	1875	1	T48669	Human EDG-2 recept	931	34	2271	1	V84632	Human secreted pro
859	34	1875	1	V33033	Dirofilaria immiti	932	34	2276	1	V84583	Human secreted pro
860	34	1875	1	V33017	Dirofilaria immiti	933	34	2294	1	V79584	Rat organic anion
861	34	1877	1	V46370	Nucleic acid encod	934	34	2298	1	X06781	Human testis secre
862	34	1883	1	V59670	Human secreted pro	935	34	2304	1	Q14176	Clone pXR2C8 encod
863	34	1890	1	V34217	Human secreted pro	936	34	2304	1	Q55374	PAR2C8 DNA. Multim
864	34	1893	1	V09852	Pig plasma membran	937	34	2304	1	T76787	Insect XR2C recept
865	34	1898	1	T79634	DNA encoding human	938	34	2304	1	T89959	D. melanogaster XR
866	34	1907	1	V84508	Human secreted pro	939	34	2307	1	V44318	Maize starch-assoc
867	34	1908	1	T73002	Human adult brain	940	34	2312	1	V21646	T. gondii chorisma
868	34	1910	1	Q50579	Asparaginylendopep	941	34	2323	1	V59524	Human secreted pro
869	34	1910	1	V27038	Human napsin B cDN	942	34	2336	1	X02889	Human secreted pro
870	34	1922	1	N70525	Sequence of cDNA i	943	34	2339	1	Q63892	Rat mACHR-6 cDNA f
871	34	1923	1	N91039	DNA sequence of ri	944	34	2339	1	T01339	Tomato acid invert
872	34	1925	1	V34271	Human secreted pro	945	34	2345	1	X18951	Human PIGR-1 encod
873	34	1931	1	Q12417	bymf gene encoding	946	34	2361	1	X04316	Human secreted pro
874	34	1934	1	T13229	OR-1 orphan recept	947	34	2377	1	Q20360	Human pro-urokinas
875	34	1946	1	V40886	Coding sequence of	948	34	2378	1	T84975	Rat glial cell der
876	34	1951	1	X37414	Human secreted pro	949	34	2400	1	Q05055	Placenta-specific
877	34	1958	1	V10328	Rat Galk2 receptor	950	34	2400	1	T9364	Nicotiana plumbagi
878	34	1961	1	T70155	BT20 protein encod	951	34	2405	1	Q70827	Protein kinase (CK
879	34	1964	1	N81586	CBT-20 cDNA encodi	952	34	2405	1	Q92962	Human HRR25-like c
880	34	1964	1	N92781	cDNA sequence enco	953	34	2424	1	V59725	Human secreted pro
881	34	1964	1	Q54354	Carcinoembryonic a	954	34	2427	1	Q04107	Human pro-urokinas
882	34	1964	1	T45064	Sequence of the hu	955	34	2430	1	Q83161	DNA encoding malic
883	34	1969	1	Q47671	Acidic leucine ami	956	34	2431	1	Q31546	Human heregulin-be
884	34	1971	1	X08428	MPD-1 cDNA. Prepar	957	34	2447	1	V54587	Human secretory pr
885	34	1972	1	Q95309	Gene encoding bovi	958	34	2460	1	V44303	Human secreted pro
886	34	1973	1	N91045	Murine programmed	959	34	2468	1	V40487	Human secreted pro
887	34	1973	1	Q55770	Human secreted pro	960	34	2469	1	V82437	Rat 25-hydroxyvita
888	34	1987	1	X27344	Achatina fulica cD	961	34	2481	1	T87940	Rat neurotrophin
889	34	1994	1	T90885	Bovine beta3 adren	962	34	2483	1	V84468	Human secreted pro
890	34	2000	1	Q74367	Human secreted pro	963	34	2487	1	X33813	Coding sequence fo
891	34	2002	1	V43605	cDNA from clone fy	964	34	2510	1	X21931	Human B-raf kinase
892	34	2002	1	V63197	Human kidney PTH/P	965	34	2519	1	V54124	Human membrane pro
893	34	2006	1	Q29607	C. felis esterase,	966	34	2524	1	T18673	Human tropinin cd
894	34	2007	1	V40754	C. felis esterase,	967	34	2549	1	N91467	Sequence of human
895	34	2007	1	V40755	Human kidney PTH/P	968	34	2550	1	V19334	Rat brain ICAM-4 p
896	34	2010	1	T15948	Human Parathyroid	969	34	2550	1	V11663	Rat brain cDNA clo
897	34	2010	1	V08391	Gene No. 20 encodi	970	34	2550	1	V34661	Rat brain cDNA clo
898	34	2017	1	V08830	Human secreted pro	971	34	2587	1	V08980	GTP-gene from Chin
899	34	2026	1	V45313	Human nuclear prol	972	34	2587	1	Q11148	Human c-IAP1. Nucl
900	34	2032	1	T42073	DNA encoding human	973	34	2589	1	T61590	Human c-IAP2. Nucl
901	34	2035	1	N70687	Rhizopus PGK1 prom	974	34	2601	1	T61591	Rat neurotrophin
902	34	2051	1	Q14944	Human metalloprote	975	34	2604	1	T73595	MACH isoform alpha
903	34	2057	1	T63123	Clone BV291_3 codi	976	34	2619	1	T61406	Human huBUB3 gene.
904	34	2060	1	V44862	Human secreted pro	977	34	2619	1	V74173	Lupin exo-(1-4)bet
905	34	2095	1	X27331	Human secreted pro	978	34	2628	1	T01014	Pig inhibitor BI-1
906	34	2096	1	T36141	Human L-asparagina	979	34	2634	1	V59067	Big transforming g
907	34	2096	1	V15795	cDNA for wild type	980	34	2639	1	V52934	Nucleotide sequenc
908	34	2104	1	V40506	Human secreted pro	981	34	2643	1	V99984	PVM051. DNA sequen
909	34	2117	1	T61258	560E1 cDNA encodin	982	34	2662	1	Q37757	

983 34 1.1 2665 1 V32592 Schwannomin-bindin
 984 34 1.1 2669 1 Q56925 Pig TGF-beta-3. Nu
 985 34 1.1 2671 1 Q03303 Entire porcine tra
 986 34 1.1 2676 1 Q02819 cDNA sequence enco
 987 34 1.1 2682 1 V63195 CDNA from clone fg
 988 34 1.1 2696 1 N60909 Plasmid pAU157 ins
 989 34 1.1 2696 1 N60889 Plasmid pAU157 seq
 990 34 1.1 2697 1 N60864 Sequence of plasm
 991 34 1.1 2744 1 V48799 Nucleotide sequenc
 992 34 1.1 2760 1 Q85425 T2 DNA-polymerase
 993 34 1.1 2760 1 Q85426 T4 DNA-polymerase
 994 34 1.1 2764 1 V63190 CDNA from clone dh
 995 34 1.1 2773 1 X07428 Homo sapiens secre
 996 34 1.1 2788 1 V55745 Human secreted pro
 997 34 1.1 2790 1 V32998 Human concentrativ
 998 34 1.1 2822 1 X52222 Protein PR0228 cDN
 999 34 1.1 2823 1 V32333 Natural killer lyl
 1000 34 1.1 2836 1 V40744 C. felis esterase,

ALIGNMENTS

RESULT 1

T99077/c
 ID T99077 standard; DNA; 45 BP.

AC T99077; 24-MAR-1998 (first entry)
 DE Probe Ena-A7 for alphaENAC coding sequence.
 KW Alpha epithelial sodium channel; alphaENACa; binding assay;
 KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
 KW salt substitute; salty taste blocker; probe; ss.
 OS Synthetic.
 OS Rattus rattus.
 PN US5693756-A.
 PD 02-DEC-1997.
 PF 23-JAN-1995; 376362.
 PR 23-JAN-1995; US-376362.
 PR 28-FEB-1994; US-202654.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI Blackshaw S, Li X, Snyder SH;
 DR WPI; 98-031814/03.
 PT Alternatively spliced epithelial sodium channel alpha subunit
 PT proteins - useful in screening assays for salty taste enhancers or
 PT blockers
 PT Disclosure; Column 6; 33pp; English.
 CC This sequence represents a probe for the coding sequence for the alpha
 CC epithelial sodium channel a (alphaENACa). AlphaENACa (see W34529) and
 CC alphaENACb (see W34530) represent the sequences of the invention. The two
 CC sodium channels are alternatively spliced forms of the
 CC amiloride-sensitive salt channel alpha subunit and can be used in
 CC membrane-transport or binding assays to identify substances that enhance
 CC or block perception of a salty taste. Enhancers could be used as salt
 CC substitutes and blockers could be used to mask salty tastes in foods and
 CC pharmaceuticals.
 SQ Sequence 45 BP; 10 A; 10 C; 12 G; 13 T;

Query Match 1.4%; Score 45; DB 1; Length 45;
 Best Local Similarity 100.0%; Pred. No. 4.5e-08;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1134 caacaatggttttcctgcactgcacagacagacagaatgactt 1178
 |
 DB 45 CAACAATGGTTGTCCTGACTGCGACAGACAGAGAAATGACTT 1

RESULT 2

T99078/c
 ID T99078 standard; DNA; 45 BP.

AC T99078; 24-MAR-1998 (first entry)
 DE Probe Ena-A9 for alphaENAC coding sequence.

KW Alpha epithelial sodium channel; alphaENACa; binding assay;
 KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
 KW salt substitute; salty taste blocker; probe; ss.
 OS Synthetic.
 OS Rattus rattus.
 PN US5693756-A.
 PD 02-DEC-1997.
 PF 23-JAN-1995; 376362.
 PR 23-JAN-1995; US-376362.
 PR 28-FEB-1994; US-202654.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI Blackshaw S, Li X, Snyder SH;
 DR WPI; 98-031814/03.
 PT Alternatively spliced epithelial sodium channel alpha subunit
 PT proteins - useful in screening assays for salty taste enhancers or
 PT blockers
 PT Disclosure; Column 7; 33pp; English.
 CC This sequence represents a probe for the coding sequence for the alpha
 CC epithelial sodium channel a (alphaENACa). AlphaENACa (see W34529) and
 CC alphaENACb (see W34530) represent the sequences of the invention. The two
 CC sodium channels are alternatively spliced forms of the
 CC amiloride-sensitive salt channel alpha subunit and can be used in
 CC membrane-transport or binding assays to identify substances that enhance
 CC or block perception of a salty taste. Enhancers could be used as salt
 CC substitutes and blockers could be used to mask salty tastes in foods and
 CC pharmaceuticals.
 SQ Sequence 45 BP; 11 A; 11 C; 10 G; 13 T;

Query Match 1.4%; Score 45; DB 1; Length 45;
 Best Local Similarity 100.0%; Pred. No. 4.5e-08;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1498 gactaccgaagcagagctctctgggctattgctattataaactg 1542
 |
 DB 45 GACTACCGAAGCAGAGCTCTCTGGGCTATTGCTATTATAAAGT 1

RESULT 3

T99072
 ID T99072 standard; DNA; 75 BP.

AC T99072; 24-MAR-1998 (first entry)
 DE Alpha epithelial sodium channel a coding sequence.
 KW Alpha epithelial sodium channel; alphaENACa; binding assay;
 KW amiloride-sensitive salt channel alpha subunit; membrane-transport;
 KW salt substitute; salty taste blocker; ss.
 OS Rattus rattus.
 PN US5693756-A.
 PD 02-DEC-1997.
 PF 23-JAN-1995; 376362.
 PR 23-JAN-1995; US-376362.
 PR 28-FEB-1994; US-202654.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI Blackshaw S, Li X, Snyder SH;
 DR WPI; 98-031814/03.
 DR P-PSDB; W34529.

PT Alternatively spliced epithelial sodium channel alpha subunit
 PT proteins - useful in screening assays for salty taste enhancers or
 PT blockers
 PS Disclosure; Column 21-22; 33pp; English.

CC This sequence represents the coding sequence for the alpha epithelial
 CC sodium channel a (alphaENACa). AlphaENACa and alphaENACb (see W34530)
 CC represent the sequences of the invention. The two sodium channels are
 CC alternatively spliced forms of the amiloride-sensitive salt channel alpha
 CC subunit and can be used in membrane-transport or binding assays to
 CC identify substances that enhance or block perception of a salty taste.
 CC Enhancers could be used as salt substitutes and blockers could be used to
 CC mask salty tastes in foods and pharmaceuticals.
 SQ Sequence 75 BP; 11 A; 22 C; 22 G; 20 T;

Query Match 1.2%; Score 38; DB 1; Length 75;

Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 tctcaagtgctggaagccttgatggtacaaactac 1614
|||||
DB 38 TCTCCAAAGTGTGGAGCCTTGATGATCAACTAC 75

RESULT 4

Q03293
ID Q03293 standard; DNA; 960 BP.

AC Q03293;

DT 03-AUG-1990 (first entry)

DE Recombinant DNA encoding human prolactin (HP)

KM Human prolactin (HP); plasmid pHP100; plasmid pDR720; plasmid pHP100.

OS Homo sapiens.

PN J02000445-A.

PD 05-JAN-1990.

PF 25-DEC-1987; 331244.

PR 25-DEC-1987; JP-331244, JP-315317.

PA (SHIK-) Shikishima Boseki K.

PI WPI; 90-047987/07.

DR P-PSDB; R05231.

PT In which promoter, Shine-Dalgarno sequence and translation

PS Initiation codon are integrated

CC Also new are bacteria (E. coli) expressing HP which contain it and the

CC prodn. of HP by their culture. A promoter, Shine-Dalgarno (SD)

CC sequence and translation initiation codon (TIC) are inserted. In

CC order, upstream of the HP gene. The HP gene is isolated from pHP100, the

CC promoter, SD sequence and TIC are obtd. from pDR720, and these are

CC ligated with a synthetic DNA linker to produce plasmid pHP100. It allows

CC large amts. of HP to be produced recombinantly.

Sequence 960 BP; 287 A; 246 C; 209 G; 218 T;

SO

Query Match 1.2%; Score 38; DB 1; Length 960;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 898 CATCAAAAAAAAAAAAAAAAAAAAAAAAAA 935

RESULT 5

V29126/C
ID V29126 standard; cDNA; 1276 BP.

AC V29126;

DT 21-AUG-1998 (first entry)

DE HPRP-1 coding sequence.

KM HPRP-1: human; disease associated protein tyrosine phosphatase; HPRP-2;

KM inflammatory cell; cancer; rheumatoid arthritis; osteoarthritis;

KW carcinoma; therapy; ss.

OS Homo sapiens.

PF Key Location/Qualifiers

FT 604..1116

FT /tag=a

FT /product= HPRP-1

PN W09814596-A2.

PD 09-APR-1998.

PF 02-OCT-1997; U17890.

PR 03-OCT-1996; US-725532.

PA (INCY-) INCYTE PHARM INC.

PI G011 SK; 98-240095/21.

DR WPI; 98-240095/21.

PT P-PSDB; W37913.

PT New isolated human protein tyrosine phosphatase(s) - are used to

PT develop products for treating e.g. inflammatory conditions such as

PT rheumatoid arthritis or osteoarthritis, or cancers

PT Claim 3; Fig 1; 58pp; English.

CC This sequence encodes the human disease associated protein tyrosine
CC phosphatase protein, designated HPRP-1, of the invention. HPRP-1 and
CC HPRP-2 are associated with inflammatory cells and various cancers. They
CC can be used to develop products for treating inflammatory conditions such
CC as rheumatoid arthritis and osteoarthritis and carcinomas of the
CC intestine, bladder, prostate, breast and brain. The products can also be
CC used for detection, diagnosis and drug screening.

Sequence 1276 BP; 407 A; 223 C; 225 G; 408 T;

Query Match 1.2%; Score 38; DB 1; Length 1276;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||

DB 41 CATCAAAAAAAAAAAAAAAAAAAAAAAAAA 4

RESULT 6

N60801
ID N60801 standard; DNA; 3557 BP.

AC N60801;

DT 24-OCT-1991 (first entry)

DE Human pro-growth hormone.

KW Pituitary gland; E.coli; ds.

OS Homo sapiens.

PN J61202689-A.

PD 08-SEP-1986.

PF 04-MAR-1985; 042404.

PR 04-MAR-1985; JP-042404.

PA (NAKA/) NAKAJIMA K.

DR WPI; 86-276387/42.

PT Plasmid for amplification of human growth hormone gene - derived

PT from human pituitary gland tissue.

PS Disclosure: Table 1-9; 9pp; Japanese.

CC Sequence is derived from human pituitary gland tissue, and allows

CC pro-growth hormone to be produced by a circular plasmid carrying an

CC amp resistance gene, from a transformed E.coli expression system in

CC an ampicillin containing medium.

See also J61202690.

Sequence 3557 BP; 927 A; 902 C; 897 G; 831 T;

SO

Query Match 1.2%; Score 38; DB 1; Length 3557;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||

DB 1149 CATCAAAAAAAAAAAAAAAAAAAAAAAAAA 1186

RESULT 7

N60847
ID N60847 standard; DNA; 3621 BP.

AC N60847;

DT 24-OCT-1991 (first entry)

DE Human pre-prolactin gene.

KW Pituitary gland; E.coli; ds.

OS Homo sapiens.

PN J61202690-A.

PD 08-SEP-1986.

PF 04-MAR-1985; 042405.

PR 04-MAR-1985; JP-042405.

PA (NAKA/) NAKAJIMA K.

DR WPI; 86-276388/42.

PT Plasmid for amplification of the human prolactin gene - derived

PT from human pituitary tissue and cultivated in ampicillin contg.

PT medium to obtain a pBR 322 ampicillin resistant gene.

PS Disclosure: Table 1-9; 9pp; Japanese.

CC Sequence is derived from human pituitary gland tissue, and allows

CC pre-prolactin to be produced by a circular plasmid carrying an amp

CC claimed full-length flea saliva protein, termed Pfispj1-113 (see
CC W82386). nfspj1-606 was isolated from a whole flea cDNA
CC library using partial clone nfspj-420 (see V73419) as probe.
CC Its nucleotide sequence shows no significant homology to known
CC sequences. Mature fispj1 protein (see W82387), and a nucleic acid
CC sequence encoding it (see V73443) are also claimed. The invention
CC is directed to methods for isolating ectoparasite saliva proteins
CC (ESPs). It provides ESps (claimed, see W82382-93), nucleic acid
CC molecules encoding them, methods for their recombinant production,
CC therapeutic compositions for treating allergic dermatitis that
CC comprise at least one ESP, as well as assay kits for testing if an
CC animal has, or is susceptible to, allergic dermatitis, and a method
CC of desensitising a host animal to allergic dermatitis using ESps.
CC The ESps can also be used for the production of antibodies useful
CC in diagnosis or in vaccines for passive immunisation against
CC allergic dermatitis.
SQ Sequence 606 BP; 173 A; 77 C; 80 G; 276 T;

Query Match 1.2%; Score 37; DB 1; length 606;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3081 attcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 44 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 8

RESULT 11
V59566
ID V59566 standard; DNA: 732 BP.
AC 06-JAN-1999 (first entry)
DT Human secreted protein gene 56 clone HSAXS65.
DE Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; lissens; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WO9839448-A2.
PD 11-SEP-1998.
PE 06-MAR-1998; U04493.
PR 02-OCT-1997; US-061060.
PR 07-MAR-1997; US-038621.
PR 07-MAR-1997; US-040161.
PR 07-MAR-1997; US-040162.
PR 07-MAR-1997; US-040163.
PR 07-MAR-1997; US-040333.
PR 07-MAR-1997; US-040334.
PR 07-MAR-1997; US-040336.
PR 07-MAR-1997; US-040626.
PR 11-APR-1997; US-043311.
PR 11-APR-1997; US-043312.
PR 11-APR-1997; US-043313.
PR 11-APR-1997; US-043314.
PR 11-APR-1997; US-043568.
PR 11-APR-1997; US-043569.
PR 11-APR-1997; US-043576.
PR 11-APR-1997; US-043580.
PR 11-APR-1997; US-043669.
PR 11-APR-1997; US-043670.
PR 11-APR-1997; US-043671.
PR 11-APR-1997; US-043672.
PR 11-APR-1997; US-043674.
PR 23-MAY-1997; US-047492.
PR 23-MAY-1997; US-047500.
PR 23-MAY-1997; US-047501.
PR 23-MAY-1997; US-047502.

PR 23-MAY-1997; US-047503.
PR 23-MAY-1997; US-047581.
PR 23-MAY-1997; US-047582.
PR 23-MAY-1997; US-047583.
PR 23-MAY-1997; US-047584.
PR 23-MAY-1997; US-047585.
PR 23-MAY-1997; US-047586.
PR 23-MAY-1997; US-047587.
PR 23-MAY-1997; US-047588.
PR 23-MAY-1997; US-047589.
PR 23-MAY-1997; US-047590.
PR 23-MAY-1997; US-047592.
PR 23-MAY-1997; US-047593.
PR 23-MAY-1997; US-047594.
PR 23-MAY-1997; US-047595.
PR 23-MAY-1997; US-047596.
PR 23-MAY-1997; US-047597.
PR 23-MAY-1997; US-047598.
PR 23-MAY-1997; US-047599.
PR 23-MAY-1997; US-047600.
PR 23-MAY-1997; US-047601.
PR 23-MAY-1997; US-047612.
PR 23-MAY-1997; US-047613.
PR 23-MAY-1997; US-047614.
PR 23-MAY-1997; US-047615.
PR 23-MAY-1997; US-047617.
PR 23-MAY-1997; US-047618.
PR 23-MAY-1997; US-047632.
PR 23-MAY-1997; US-047633.
PR 06-JUN-1997; US-048964.
PR 06-JUN-1997; US-048974.
PR 13-JUN-1997; US-049610.
PR 08-JUL-1997; US-051926.
PR 16-JUL-1997; US-052874.
PR 18-AUG-1997; US-055724.
PR 22-AUG-1997; US-056630.
PR 22-AUG-1997; US-056631.
PR 22-AUG-1997; US-056632.
PR 22-AUG-1997; US-056636.
PR 22-AUG-1997; US-056637.
PR 22-AUG-1997; US-056662.
PR 22-AUG-1997; US-056664.
PR 22-AUG-1997; US-056845.
PR 22-AUG-1997; US-056862.
PR 22-AUG-1997; US-056864.
PR 22-AUG-1997; US-056872.
PR 22-AUG-1997; US-056874.
PR 22-AUG-1997; US-056875.
PR 22-AUG-1997; US-056876.
PR 22-AUG-1997; US-056877.
PR 22-AUG-1997; US-056878.
PR 22-AUG-1997; US-056879.
PR 22-AUG-1997; US-056880.
PR 22-AUG-1997; US-056881.
PR 22-AUG-1997; US-056882.
PR 22-AUG-1997; US-056884.
PR 22-AUG-1997; US-056886.
PR 22-AUG-1997; US-056887.
PR 22-AUG-1997; US-056888.
PR 22-AUG-1997; US-056889.
PR 22-AUG-1997; US-056892.
PR 22-AUG-1997; US-056893.
PR 22-AUG-1997; US-056894.
PR 22-AUG-1997; US-056903.
PR 22-AUG-1997; US-056908.
PR 22-AUG-1997; US-056909.
PR 22-AUG-1997; US-056910.
PR 22-AUG-1997; US-056911.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057659.
PR 05-SEP-1997; US-057661.
PR 12-SEP-1997; US-057685.
PA (HUMA-) HUMAN GENOME SCI INC.

PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA,
PI Feng P, Ferlie AM, Fischer CL, Florence KA, Greene JM, Hu JS,
PI Kyaw H, Lafleur DM, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA,
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
DR WPI: 98-506364/43.
DR P-PSDB: W74786.
PT New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1; Page 289-290; 721pp; English.
CC This sequence represents a nucleic acid molecule designated Gene 56 from
CC the human cDNA clone HSAS65 (deposited as clone ATCC 97899 and ATCC
CC 209045) which encodes a secreted human protein. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. V59502) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 186 polynucleotides, based on
CC which tissues they are most highly expressed in (see V59511 for described
CC uses).
SQ Sequence 732 BP; 220 A; 167 C; 183 G; 162 T;

Query Match 1.2%; Score 37; DB 1; Length 732;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 674 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 710

RESULT 12
ID N50414 standard; DNA; 970 BP.
AC N50414;
DT 15-JUN-1992 (first entry)
DE Sequence encoding pig elastase I.
KW Arteriosclerosis; E.coli; ds.
OS Sus scrofa domestica.
FH Key Location/Qualifiers
FT cds 1..801
FT /+tag- a
FN J60186284-A.
PD 21-SEP-1985.
PF 07-MAR-1984; 043155.
PR 07-MAR-1984; JP-043155.
PA (KIRI) KIRIN BREWERY KK.
DR WPI: 85-273645/44.
DR P-PSDB: P50360.
PT Biochemical prepn. of pig-elastase I - uses DNA chain to produce
PT plasmid to transform Escherichia coli or Saccharomyces
PT microorganism.
PS Claim 1; Fig 1; 16pp; Japanese.
CC The gene product may be expressed from an E.coli or Saccharomyces
CC host transformed with a plasmid encoding the pig elastase I gene.
CC Elastase I is useful in the treatment of arteriosclerosis, and may be
CC efficiently produced by this technique.
SQ Sequence 970 BP; 247 A; 261 C; 262 G; 200 T;

Query Match 1.2%; Score 37; DB 1; Length 970;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 932 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 968

RESULT 13
ID T73851 standard; cDNA; 1001 BP.
AC T73851;
DT 08-SEP-1997 (first entry)
DE Full length 3' UTR of prohibitin wild-type B allele.
KW Prohibitin; tumor suppressor gene; loss of heterozygosity; LOH;
KW cell immortalisation; complementation group B; mutation; 3' UTR;
KW untranslated region; diagnosis; susceptibility; cancer; breast;
KW screen; genotype; treatment; early stage tumour; ds.
OS Synthetic.
PN WO9640919-A1.
PD 19-DEC-1996.
PF 07-JUN-1996; U09687.
PR 07-JUN-1995; US-473486.
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
PI Dellorco RT, Jupe ER, King RL, Liu X, McClung JK;
DR WPI: 97-052321/05.
PT Prohibitin gene 3' untranslated region - reintroduction of which
PT into early stage tumours is useful in cancer treatment
PS Example 1; Page 41-42; 60pp; English.

CC Prohibitin, an evolutionary conserved gene that possesses anti-
CC proliferative activity, is postulated to be a tumour suppressor gene
CC whose expression, when lost, contributes to the immortalisation of cells
CC from one or more of the four complementation groups proposed by Pereira
CC and Smith. The human prohibitin gene has been localised to chromosome 17q21
CC and is apparently involved in the process of immortalisation in a
CC group of human cells that are classified in complementation group B.
CC Mutations in the 3' untranslated region (3' UTR) of the B type allele are
CC diagnostic for increased susceptibility to cancer, particularly breast
CC cancer. Diagnosis comprises determining an individual's prohibitin
CC genotype and then screening for alterations in the 3' UTR of the
CC prohibitin gene. Also cancer can be treated by reintroducing at least a
CC portion of the wild type 3' UTR of a prohibitin gene into early stage
CC tumours. The present sequence is that of the full 3' UTR. The cDNA and
CC genomic sequences are identical, showing that the full 3' UTR is
CC transcribed from the genome without introns.
SQ Sequence 1001 BP; 267 A; 253 C; 250 G; 231 T;

Query Match 1.2%; Score 37; DB 1; Length 1001;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 940 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 976

RESULT 14
ID V69433 standard; cDNA; 1889 BP.
AC V69433;
DT 15-MAR-1999 (first entry)
DE Human edg-6 cDNA #2.
KW Endothelial differentiation gene receptor homologue; human; HEDG;
KW EDG-6; detection; diagnosis; drug screening; treatment; inflammation;
KW disease; viral; bacterial; fungal infection; allergic response; injury;
KW hereditary disease; lymphoma; carcinoma; lymphoid; neuronal; cascade;
KW lymphocyte trafficking; leukocyte trafficking; signal transduction;
KW rheumatoid synovium; autoimmune disorder; ss.
OS Homo sapiens.
PN WO9853062-A1.
PD 26-NOV-1998.
PF 22-MAY-1998; CA0487.
PR 22-MAY-1997; US-861747.
PA (ALIX) ALTELIX BIOPHARMACEUTICALS INC.
PI Munroe DG, Vyas TB;
DR WPI: 99-070147/06.
PT New isolated human endothelial differentiation gene-6 receptor
PT homologue - used to develop products for treating e.g. infections,

PT allergic responses, trauma, hereditary diseases, lymphoma, carcinoma
PT or autoimmune disorders
PS Disclosure: Fig 2, 59pp. English.
CC This sequence encodes a novel human endothelial differentiation gene
CC (EDG)-6 receptor homologue designated HEDG. The encoded protein can be
CC used for detection, diagnosis and drug screening. The protein can also
CC be used to treat inflammation or diseases including viral, bacterial, or
CC fungal infections, allergic responses, mechanical injury associated with
CC trauma, hereditary diseases, lymphoma or carcinoma, or other conditions
CC which activate the genes of lymphoid or neuronal tissues. It can also be
CC used for treating problems involving excessive lymphocyte and leukocyte
CC trafficking. Inhibitors of HEDG are useful for controlling signal
CC transduction and signalling cascades in cells of the rheumatoid synovium.
CC The protein can also be used for treating autoimmune disorders such as
CC myasthenia gravis.
SQ Sequence 1889 BP; 360 A; 575 C; 537 G; 416 T;

Query Match 1.2%; Score 37; DB 1; Length 1889;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 1845 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 1881

RESULT 15
Q25156
ID Q25156 standard; cDNA; 2158 BP.
AC Q25156;
DE 18-NOV-1992 (first entry)
DE Alpha-GalNAC from pAGB-3.
KW Lysosome; Schindler disease; infantile neuroaxonal dystrophy; ss.
OS Homo sapiens.
FH Key location/Qualifiers
FT cds 345..1580
FT /*tag= a
FT /label= alpha-GalNAC
FT signal_peptide 345..395
FT /*tag= b
FT mat_peptide 396..1580
FT /*tag= c
FT poly_a_signal 2073..2078
FT /*tag= d
FT protein_bind 2025..2029
FT /*tag= e
FT /note= "recognised by the U4 small nuclear
FT ribonucleoprotein"
FN W09207936-A.
PD 14-MAY-1992.
PE 23-OCT-1991; U07872.
PR 24-OCT-1990; US-602608.
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE.
PI Bishop DE, Desnick RJ, Ioannou YA, Wang AM;
DR P-PSDB; R24291.
DE Cloning and expression of alpha-n-acetyl-galactose aminidase -
PT used in enzyme replacement therapy for Schindler disease
PS Disclosure: Fig 2 (A-D); 71pp. English.
CC The sequence is of the pAGB-3 cDNA insert contg. the complete coding
CC region for human alpha-GalNAC.
CC The availability of the full length cDNA for alpha-GalNAC allows
CC the study of the genomic organisation and evolution of this
CC lysosomal gene, and the characterisation of molecular lesions
CC causing Schindler disease.
SQ Sequence 2158 BP; 517 A; 610 C; 576 G; 455 T;

Query Match 1.2%; Score 37; DB 1; Length 2158;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3081 atcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 2090 ATCAAAAAAAAAAAAAAAAAAAAAAAAAA 2126

RESULT 16
ID V55872 standard; cDNA; 2702 BP.
AC V55872;
DE 18-NOV-1998 (first entry)
DE Plant acetylactate synthase (ALS) large subunit protein encoding cDNA.
KW ALS; small subunit; acetylactate synthase; plant; ssuALS; lsuALS;
KW Nicotiana glumabaginfolia; plasmid vector; herbicide; holoenzyme;
KW large subunit; ss.
OS Nicotiana glumabaginfolia.
FH Key location/Qualifiers
FT cds 492..2492
FT /*tag= a
FT /product= "ALS large subunit protein"
FN W09837206-A1.
PD 27-AUG-1998.
PE 23-FEB-1998; U03506.
PR 24-FEB-1997; US-039148.
PA (DUPO) DU PONT DE NEMOURS & CO E. I.
PI Abell LM, Hershey HP;
DR WPI; 98-467568/40.
DE P-PSDB; W79141.
DE New nucleic acid encoding a plant aceto:lactate synthase small
PT sub:unit - that combines with the large sub:unit to give a
PT holoenzyme having higher activity than the large sub:unit alone
PS Example 6; Pages 36-39; 47pp. English.
CC This cDNA encodes the large subunit of a plant acetylactate synthase.
CC The invention provides a cDNA sequence contained in the plasmid vector comprising
CC encoding the small subunit of ALS (ssuALS). The plasmid vector comprising
CC the ssuALS nucleotide operably linked to a regulatory sequence can be
CC used to transform host cells for the recombinant production of the ssuALS
CC protein which is used for evaluating a compound for acetylactate synthase
CC inhibition and so for selecting potential herbicides. Evaluation of a
CC compound for acetylactate synthase (ALS) inhibition comprises expression
CC and purification of plant ssuALS from the transformed host, mixing this
CC ssuALS with the large subunit of ALS to form a holoenzyme which is then
CC treated with a test compound. Treated and untreated holoenzyme activity
CC are compared to select compounds with potential for herbicidal activity.
CC Previously, large subunits of plant ALS (lsuALS) have been isolated,
CC but the existence of ssuALS had not been verified. Mixing the two
CC subunits results in a holoenzyme that has a 4-15 fold increase in
CC specific activity over lsuALS alone.
SQ Sequence 2702 BP; 668 A; 645 C; 642 G; 747 T;

Query Match 1.2%; Score 37; DB 1; Length 2702;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3080 catcaaaaaaaaaaaaaaaaaaaaaaaaaa 3116
Db 2660 CATCAAAAAAAAAAAAAAAAAAAAAAAAAA 2696

RESULT 17
ID X51737 standard; DNA; 340 BP.
AC X51737;
DE 17-JUN-1999 (first entry)
DE DNA encoding a human secreted protein.
KW Human secreted protein; cancer; immune disorder; infection;
KW inflammatory disorder; skin disorder; tumour; atherosclerosis;
KW restenosis; autoimmune disorder; Alzheimer's disease;
KW peripheral neuropathy; trauma; spinal cord injury; allergy;
KW hemopoietic disorder; skeletal disorder; neurological disorder;
KW arthritic disorder; asthma; immunodeficiency disease; AIDS;
KW transplant rejection; ss.
OS Homo sapiens.

ID X22245 standard; DNA; 437 BP.
AC X22245:
DE 18-MAY-1999 (first entry)
KW Human secreted protein gene 35 clone HMADU73.
KW tumour; secreted protein; gene therapy; protein therapy; cancer; weight;
KW tumour; chromosome mapping; forensic; haematological disease; allergy;
KW inflammation; cell proliferation; viral infection; wound healing;
KW modulation; appetite; behaviour; food additive; preservative; ss.
OS Homo sapiens.
PN WO903990-A1.
PD 28-JAN-1999.
PE 15-JUL-1998: U14613.
PR 18-AUG-1997: US-056361.
PR 16-JUL-1997: US-052661.
PR 16-JUL-1997: US-052870.
PR 16-JUL-1997: US-052871.
PR 16-JUL-1997: US-052872.
PR 16-JUL-1997: US-052873.
PR 16-JUL-1997: US-052874.
PR 16-JUL-1997: US-052875.
PR 22-JUL-1997: US-053440.
PR 22-JUL-1997: US-053441.
PR 22-JUL-1997: US-053442.
PR 18-AUG-1997: US-055683.
PR 18-AUG-1997: US-055724.
PR 18-AUG-1997: US-055725.
PR 18-AUG-1997: US-055726.
PR 18-AUG-1997: US-055946.
PR 18-AUG-1997: US-055952.
PR 18-AUG-1997: US-055985.
PR 18-AUG-1997: US-055989.
PR 18-AUG-1997: US-056359.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Duan R, Feng P, Ferlie AM, Florence KA, Fouad J,
PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PE,
PI Yu G;
DR WPI: 99-132234/11.
DR P-PSDB: Y01417.
PT New nucleic acids encoding secreted human proteins - potentially
PT useful for treating and diagnosing diseases and identifying specific
PT binding agents
PS Claim 4: Page 187: 251pp: English.
CC The invention relates to nucleic acid sequences (X22211 to X22283)
CC encoding human secreted proteins (Y01383 to Y01454). The secreted protein
CC gene sequences are deposited with the ATCC under deposit number ATCC
CC 209138, 209139 or 209141. Host cells containing vectors comprising the
CC nucleic acid sequences are used for the recombinant expression of the
CC secreted proteins. The polynucleotide and amino acid sequences are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. Pathological conditions can be also diagnosed by
CC determining the amount of the new polypeptides in a sample or by the
CC presence of mutations in the new polynucleotides. The nucleic acid
CC sequences, or its fragments, are useful for chromosome identification and
CC mapping; as antisense and triplex-forming therapeutics; in gene therapy;
CC for (forensic) identification of individuals; as molecular weight
CC markers; to identify related sequences or specific mRNA; in preparation
CC of oligomers and to raise anti-DNA antibodies. Antibodies are useful as
CC immunosay reagents (including for in vivo imaging) and therapeutically
CC to inhibit or activate particular polypeptides. A very wide range of
CC disorders may be treated with the polynucleotide and polypeptide
CC sequences, e.g. autoimmune or haematological diseases, allergy,
CC inflammation, cancer or other forms of cell proliferation, viral or other
CC infections. The sequences may also be useful in wound healing, to
CC modulate differentiation of embryonic stem cells, to modulate weight,
CC appetite, behaviour etc. and as food additive or preservative. The
CC present sequence represents a gene encoding a human secreted protein
CC (see descriptor line for gene number and clone identification).
SQ Sequence 437 BP; 136 A; 128 C; 97 G; 72 T;

Query Match 1.2%; Score 36; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 380 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 415
RESULT 21
X04408
ID X04408 standard; DNA; 505 BP.
AC X04408:
DE 13-APR-1999 (first entry)
KW Human secreted protein gene.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WO9856804-A1.
PD 17-DEC-1998.
PE 11-JUN-1998: U12125.
PR 02-OCT-1997: US-061060.
PR 13-JUN-1997: US-049547.
PR 13-JUN-1997: US-049548.
PR 13-JUN-1997: US-049549.
PR 13-JUN-1997: US-049550.
PR 13-JUN-1997: US-049606.
PR 13-JUN-1997: US-049607.
PR 13-JUN-1997: US-049608.
PR 13-JUN-1997: US-049609.
PR 13-JUN-1997: US-049610.
PR 13-JUN-1997: US-049611.
PR 13-JUN-1997: US-050566.
PR 13-JUN-1997: US-050901.
PR 13-JUN-1997: US-052989.
PR 08-JUL-1997: US-051919.
PR 18-AUG-1997: US-055984.
PR 12-SEP-1997: US-058665.
PR 12-SEP-1997: US-058666.
PR 12-SEP-1997: US-058668.
PR 12-SEP-1997: US-058750.
PR 12-SEP-1997: US-058971.
PR 12-SEP-1997: US-058972.
PR 12-SEP-1997: US-058975.
PR 02-OCT-1997: US-060834.
PR 02-OCT-1997: US-060841.
PR 02-OCT-1997: US-060844.
PR 02-OCT-1997: US-060865.
PR 02-OCT-1997: US-061059.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Ferlie AM, Feng P, Greene JM, Lafleur DM,
PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P,
PI Yu G;
DR WPI: 99-080881/07.
DR P-PSDB: W78223.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1: Page 262: 380pp: English.
CC This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number, and the clone it is derived from, are
CC detailed in the descriptor line. The gene can be used to generate fusion
CC proteins by linking to the gene to a human immunoglobulin Fc portion
CC (e.g. X04302) for increasing the stability of the fused protein as
CC compared to the human protein only.
CC The invention relates to 86 novel genes and their fragments (nucleic acid
CC sequences: X04311-X04410; amino acid sequences W78126-W78225) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample

CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 86 polynucleotides, based on
CC which tissues they are most highly expressed in (see X04311 for described
CC uses).
SQ Sequence 505 BP; 135 A; 150 C; 120 G; 97 T;

Query Match 1.2%; Score 36; DB 1; Length 505;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 458 TCAAAAAAAAAAAAAAAAAAAAAA 493

RESULT 22

T64563
ID T64563 standard; cDNA to mRNA; 558 BP.
AC T64563;
DT 01-DEC-1997 (first entry)
DE Mouse thymus and activation regulated cytokine (TARC) cDNA.
KM Mouse; thymus; activation; regulation; chemokine; TARC; CC type;
KM peripheral; blood; monocyte; CCR4; T cell; lymphocyte;
OS Immunomodulation; inflammation; murine; ds.
FH Key Location/Qualifiers
FT cds 2..283
FT /*tag= a
PN MO9711969-A1.
PD 03-APR-1997.
PE 27-SEP-1996; J02801.
PR 13-MAR-1996; JP-056044.
PR 27-SEP-1995; JP-249457.
PA (SHIO) SHIONOGI & CO LTD.
PI Imai T, Yoshida T, Yoshie O;
DR WPI; 97-212853/19.
DR P-PSDB; W14918.
PT New CC-chemokine-like protein - is expressed by peripheral blood
PT monocytes under immunological stimulus, useful as immunomodulator
PT and antiinflammatory agent
PS Claim 12; Pages 59-60; 102pp; Japanese.
CC The present sequence encodes mouse thymus and activation regulated
CC chemokine (TARC), which is similar to CC type chemokines and
CC expressed in peripheral blood monocytes under immunological
CC stimulus, e.g. phyto-haemagglutinin or soluble cytokine, and the
CC thymus, but not by the spleen. TARC is active against CCR4
CC expressing cells, e.g. T cells, peripheral lymphocytes and
CC activated peripheral T cells, especially Hut 78, Hut 102, Jurkat,
CC MT2 and MT4 cells, and is useful as an immunomodulator or
CC anti-inflammatory agent.
SQ Sequence 558 BP; 210 A; 124 C; 121 G; 103 T;

Query Match 1.2%; Score 36; DB 1; Length 558;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 464 TCAAAAAAAAAAAAAAAAAAAAAA 499

RESULT 23

T69172
ID T69172 standard; cDNA; 581 BP.
AC T69172;
DT 12-JAN-1998 (first entry)
DE Trypanosoma cruzi antigen (ubiquitin) cDNA.
KM Antigen; epitope; vaccine; protective immunity; Chagas disease;
KM diagnosis; therapy; immunoassay; ubiquitin; ss.
OS Trypanosoma cruzi Tulane strain C2.
FH Key Location/Qualifiers

FT CDS 45..431
FT /*tag= a
PN MO9718475-A1.
PD 22-NOV-1997.
PE 14-NOV-1996; U18624.
PR 14-NOV-1995; US-557309.
PA (CORI-) CORIXA CORP.
PI Houghton RL, Lodes MJ, Reed SG, Skeiky YAM;
DR WPI; 97-289413/26.
DR P-PSDB; W26534.

PT Diagnosing Trypanosoma cruzi infection by detecting antibodies to
PT novel antigens - which are useful in vaccines to provide protective
PT immunity against Chagas' disease
PS Claim 1; Page 50-51; 110pp; English.
CC This novel DNA molecule was isolated by screening a Trypanosoma
CC cruzi cDNA expression library with pools of sera from infected
CC individuals. It encodes T. cruzi ubiquitin (see W26534). 22
CC isolated genomic DNA of cDNA molecules (T69151-72) encode full-
CC length antigens (see W26530-41), or epitope-containing repeat
CC sequences (see W19094-102, W19079-86 and W26542-44) of native
CC antigens, that can be used in a variety of immunoassays for
CC detecting T. cruzi infection in a blood, serum, plasma, saliva,
CC cerebrospinal fluid or urine sample. The polypeptides are also
CC useful in vaccines and pharmaceutical compositions for inducing
CC protective immunity against Chagas disease. The isolated DNA
CC molecules can be used for recombinant production of the antigenic
CC polypeptides.
SQ Sequence 581 BP; 160 A; 139 C; 155 G; 127 T;

Query Match 1.2%; Score 36; DB 1; Length 581;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 543 TCAAAAAAAAAAAAAAAAAAAAAA 578

RESULT 24

T64562
ID T64562 standard; cDNA to mRNA; 582 BP.
AC T64562;
DT 01-DEC-1997 (first entry)
DE Human thymus and activation regulated cytokine (TARC) cDNA.
KM Human; thymus; activation; regulation; chemokine; TARC; CC type;
KM peripheral; blood; monocyte; CCR4; T cell; lymphocyte;
OS Immunomodulation; inflammation; ds.
FH Key Location/Qualifiers
FT cds 53..337
FT /*tag= a
PN MO9711969-A1.
PD 03-APR-1997.
PE 27-SEP-1996; J02801.
PR 13-MAR-1996; JP-056044.
PR 27-SEP-1995; JP-249457.
PA (SHIO) SHIONOGI & CO LTD.
PI Imai T, Yoshida T, Yoshie O;
DR WPI; 97-212853/19.
DR P-PSDB; W14917.
PT New CC-chemokine-like protein - is expressed by peripheral blood
PT monocytes under immunological stimulus, useful as immunomodulator
PT and antiinflammatory agent
PS Claim 10; Pages 58-59; 102pp; Japanese.
CC The present sequence encodes human thymus and activation regulated
CC chemokine (TARC), which is similar to CC type chemokines and
CC expressed in peripheral blood monocytes under immunological
CC stimulus, e.g. phyto-haemagglutinin or soluble cytokine, and the
CC thymus, but not by the spleen. TARC is active against CCR4
CC expressing cells, e.g. T cells, peripheral lymphocytes and
CC activated peripheral T cells, especially Hut 78, Hut 102, Jurkat,
CC MT2 and MT4 cells, and is useful as an immunomodulator or

CC anti-inflammatory agent.
SQ Sequence 582 BP; 162 A; 168 C; 149 G; 103 T;
Query Match 1.2%; Score 36; DB 1; Length 582;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 537 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 572
RESULT 25
ID X20419 standard; DNA; 602 BP.
AC X20419;
DE Human secreted protein gene 8.
KW Human secreted protein; fusion protein; gene therapy; protein therapy;
diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
KW Homo sapiens.
OS Homo sapiens.
PN MO9906423-A1.
PD 11-FEB-1999.
PF 29-JUL-1998; U15949.
PR 19-AUG-1997; US-056730.
PR 30-JUL-1997; US-054209.
PR 30-JUL-1997; US-054211.
PR 30-JUL-1997; US-054212.
PR 30-JUL-1997; US-054213.
PR 30-JUL-1997; US-054214.
PR 30-JUL-1997; US-054215.
PR 30-JUL-1997; US-054217.
PR 30-JUL-1997; US-054218.
PR 30-JUL-1997; US-054234.
PR 30-JUL-1997; US-054236.
PR 18-AUG-1997; US-055968.
PR 18-AUG-1997; US-055969.
PR 18-AUG-1997; US-055972.
PR 19-AUG-1997; US-056534.
PR 19-AUG-1997; US-056534.
PR 19-AUG-1997; US-056554.
PR 19-AUG-1997; US-056561.
PR 19-AUG-1997; US-056727.
PR 19-AUG-1997; US-056729.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Carter KC, Endress GA, Fan P, Feng P, Kyaw H, Lafleur DM,
Li Y, Moore PA, Rosen CA, Ruben SM, Shi Y, Wei Y,
Zeng Z,
WPI; 99-1363691/13.
DR P-PSDB; Y00265.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1; Page 214; 312pp; English.
CC This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number is given in the descriptor line. The gene
CC can be used to generate fusion proteins by linking to the gene to a human
CC immunoglobulin Fc portion (e.g. X20403) for increasing the stability of
CC the fused protein as compared to the human protein only.
CC The invention relates to 83 novel genes and their fragments (nucleic acid
CC sequences: X20412-X20499; amino acid sequences Y00258-Y00377) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 86 polynucleotides, based on

CC which tissues they are most highly expressed in (see X20412 for described
CC uses).
SQ Sequence 602 BP; 194 A; 133 C; 152 G; 123 T;
Query Match 1.2%; Score 36; DB 1; Length 602;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 528 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 563
RESULT 26
ID Q72956 standard; CDNA; 684 BP.
AC Q72956;
DE Clone pSP65-Xai-11 (first entry)
KW Clone pSP65-Xai-11 encoding recombinant Factor Xa inhibitor (FXai).
KW Leech, factor Xa inhibitor; ss.
OS Hirudo medicinalis.
FH Key Location/Qualifiers
FT cds 139..684
FT misc_difference 1..684
FT /*tag- a
FT /*tag- b
FT /*tag- N= unspecified base
PN MO9423709-A.
PD 27-OCT-1994.
PF 08-APR-1994; U03918.
PR 09-APR-1993; US-045804.
PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
PA (YISS) YISSUM RES & DEV CO.
PI Fischer M, Goldlust A, Guy R, Levanon A, Panet A,
PI Rigbi M, Werber MM, Zeelon EP;
DR WPI; 94-341657/42.
DR P-PSDB; R62619.
PT Recombinant factor 10a inhibitor of Hirudo medicinalis - for
PT treating excessive blood coagulation, partic. thrombosis, also
PT related DNA, vectors, transformed cells and antibodies
PS Claim 3; Figure 10; 107pp; English.
CC Total RNA was extracted from 120 leeches and used to prepare a
CC cDNA library which was screened for clones contg. FXai DNA using
CC as probe radiolabeled DNA from the PCR derived plasmid pSP65-Xai-
CC 4 (see Q72957). The plasmid of one of the positive clones - pSK-
CC Xai-11 (clone 11) contains a 684 bp insert (see Q72956) essentially
CC identical to the previously obt. PCR derived sequences but contg.
CC 78 additional bps at its 5' end. It encodes a prepeptide of 110
CC AAs (incl. the N-terminal Met) extending from Met 1 to Gly 110 (see
CC R62619). AAs 1-25 constitute a leader sequence, and the mature
CC protein comprises Tyr 26 to Gly 110. The sequence Tyr 2 - Gly 86
CC of clone 4 (Q72957) is equiv. to the sequence Tyr 26 - Gly 110
CC of clone 11 (Q72956). There are only two differences between
CC clones 4 and 11. The polypeptide expressed by clone 11 is
CC referred to as recombinant FXai. A plasmid comprising the DNA
CC designated pSP65-Xai-11 and deposited under ATCC Accession No.
CC 69138 is claimed.
SQ Sequence 684 BP; 263 A; 110 C; 120 G; 190 T;
Query Match 1.2%; Score 36; DB 1; Length 684;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
Db 636 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 671
RESULT 27
ID X30351 standard; DNA; 687 BP.

AC X30351;
DT 14-MAY-1999 (first entry)
DE DNA encoding a human secreted protein.
KW Secreted protein; cancer; neurodegenerative disorder;
KW developmental abnormality; foetal deficiency; blood disorder;
KW CNS disorder; immune system disease; autoimmune disease; hepatic disease;
KW renal disease; diabetes; inflammation; allergy; ischemic shock;
KW Alzheimer's; cognitive disorder; schizophrenia; cardiovascular disorder;
KW prostate disease; asthma; osteoporosis; arthritis; ss.
OS Homo sapiens.
PN M09907891-A1.
PD 18-FEB-1999.
PF 04-AUG-1998; U16235.
PR 19-AUG-1997; US-056732.
PR 05-AUG-1997; US-054798.
PR 05-AUG-1997; US-054803.
PR 05-AUG-1997; US-054804.
PR 05-AUG-1997; US-054806.
PR 05-AUG-1997; US-054807.
PR 05-AUG-1997; US-054808.
PR 05-AUG-1997; US-054809.
PR 05-AUG-1997; US-055309.
PR 05-AUG-1997; US-055310.
PR 05-AUG-1997; US-055312.
PR 05-AUG-1997; US-055386.
PR 05-AUG-1997; US-055311.
PR 18-AUG-1997; US-055970.
PR 18-AUG-1997; US-055986.
PR 19-AUG-1997; US-056365.
PR 19-AUG-1997; US-056366.
PR 19-AUG-1997; US-056557.
PR 19-AUG-1997; US-056370.
PR 19-AUG-1997; US-056371.
PR 19-AUG-1997; US-056563.
PR 19-AUG-1997; US-056731.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Ferrie AM, Greene JM, Janat F, Ni J,
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR, Young PE, Yu G;
DR WPI; 99-167452/14.
DR P-PSDB; Y10830.
PT New isolated human genes encoding secreted polypeptides - useful for
PT diagnosis and treatment of pathological diseases
PS Claim 3; Page 247; 331pp; English.
CC The specification describes secreted proteins and their corresponding
CC polynucleotides which are useful for preventing, treating or ameliorating
CC medical conditions, e.g. by protein or gene therapy. Pathological
CC conditions can also be diagnosed by determining the amount of the
CC secreted polypeptides in a sample or by determining the presence of
CC mutations in the polynucleotides. Specific uses are described for each
CC of the products, based on which tissues they are most highly
CC expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumours, neurodegenerative disorders, developmental
CC abnormalities and foetal deficiencies, blood disorders, CNS disorders,
CC diseases of the immune system, autoimmune diseases, hepatic disorders,
CC disease, diabetes, inflammation, allergies, ischemic shock, Alzheimer's
CC and cognitive disorders, schizophrenia, cardiovascular disorders,
CC prostate diseases, asthma, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, diseases of testes, lung or
CC thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
SQ Sequence 687 BP; 236 A; 123 C; 152 G; 173 T;

Query Match 1.2%; Score 36; DB 1; Length 687;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 634 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 669

RESULT 28

Q22005
ID Q22005 standard; DNA; 742 BP.
AC Q22005;
DT 09-JUN-1992 (first entry)
DE Sequence encoding Cor a I allergen of hazel.
KW Fagales; hazel; IgE; tree allergen; pollen; antibodies; T-cell
KW response; ss.
OS Fagales hazel.
FH Key cds Location/Qualifiers
FT 1..483
FT /*tag= a
FT /*product= Aln_g-I
FT 692..742
FT /*tag= b
FT poly_a_site
FT
FT W09202621-A.
PN 20-FEB-1992.
PD 06-MAY-1991; E01479.
PF 08-AUG-1990; AT-001688.
PR 11-APR-1991; US-683831.
PR (BIOM-) BIOMAY BIOECHAN PRO.
PI Breitenbach H, Reikertorfer A, Valenta R, Hoffmann-Sommergruber K;
PI WPI; 92-080075/10.
DR P-PSDB; R21793.
DR Aln g I, Cor a I and Bet v I allergens - and DNA from alder,
PT hazel and birch, useful in diagnosis or therapy of allergic
PT diseases
PS Claim 8; Page 28; 54pp; English.
CC The cDNA sequence encoding the Cor a I allergen of hazel was obt
CC from the polyA+ mRNA isolated from ripe hazel pollen. The resulting
CC cDNA was amplified by PCR using primers whose sequences were derived
CC from the N-terminal amino-acid sequence of Cor a I. The DNA fragment
CC produced was cloned into the pBluescript KS vector and transformed
CC into E. coli XL1-Blue cells. The sequence allows mammals to be
CC tested for allergic reactions to specific tree allergens. The
CC derived polypeptide may be used to challenge the mammal to elicit
CC bronchial, conjunctival, dermal, nasal or oral provocation. The
CC polypeptide may be used to treat a mammal afflicted with a pollen
CC allergy. It is administered in an amount sufficient to hyposenstise
CC the mammal to Cor a I.
CC See also Q22001-8 and Q21974.
SQ Sequence 742 BP; 250 A; 134 C; 172 G; 186 T;

Query Match 1.2%; Score 36; DB 1; Length 742;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 691 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 726

RESULT 29
ID T59912 standard; cDNA; 780 BP.
AC T59912;
DT 04-JUN-1997 (first entry)
DE Human transcription regulatory factor SSB7 cDNA.
KW Transcription regulatory factor; suppressor of RNA polymerase B;
KW SSB7; RNA polymerase II; holoenzyme; SWI/SNF; ss.
OS Homo sapiens.
FH Key cds Location/Qualifiers
FT 10..444
FT /*tag= a
FT 627
FT allele
FT /*tag= b
FT /*note= "there is a polymorphism at position 627;
FT in some individuals it is GATC (a
FT restriction site for Sau3A), in others it
FT is GATT"
PN W09708301-A1.
PD 06-MAR-1997.

```
PF 28-AUG-1996: U14192.
PR 31-AUG-1995: US-521872.
PR 11-OCT-1995: US-540804.
PR 26-JAN-1996: US-590399.
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI Chao DM, Koleske AJ, Thompson CM, Young RA;
DR WPI: 97-179258/16.
DR P-PSDB: W13829.
PT Purified RNA polymerase II holo:enzyme - comprises RNA polymerase II
PT and one or more regulatory proteins, pref. suppressor of RNA
PT polymerase B proteins or SWI/SNF proteins
PS Claim 11: Fig 17a: 154pp: English.
CC A cDNA clone (T59912) codes for the human homologue (W13829),
CC hSRB7, of the yeast transcription regulatory factor YSRB7 (see
CC also W13824). It was cloned and sequenced using 3 overlapping
CC expressed sequence tags. hSRB7 is 35% identical to YSRB7,
CC complements a YSRB7 deletion and, like its yeast counterpart, binds
CC to the C-terminal domain of RNA polymerase II. It forms part of a
CC holoenzyme complex that supports active transcription. SRB
CC nucleic acids (see also T59904-11) and RNA polymerase II holoenzymes
CC can be used to treat diseases resulting from alteration or deletion
CC of the SRB gene, pref. by gene transfer technology. Probes based on
CC the genes can be used to detect SRB nucleic acids.
SQ Sequence 780 BP; 289 A; 145 C; 141 G; 205 T;

Query Match 1.2%; Score 36; DB 1; Length 780;
Best Local Similarity 100.0%; Pred. No. 6; 7e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 729 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 764

RESULT 30
N92416
ID N92416 standard; cDNA: 794 BP.
AC N92416;
DT 15-MAR-1992 (first entry)
DE Sequence of the S3 cDNA clone.
KM Self-incompatible plant; S-gene: S-protein; ss.
OS Nicotiana glauca.
PN EP-343947-R.
PD 29-NOV-1989.
PR 24-MAY-1989; 305233.
PR 29-OCT-1985; US-792435.
PR 21-APR-1986; US-854139.
PR 25-MAY-1988; US-198781.
PA (LUBR ) LUBRIZOL GENETICS I.
PI Clarke AE, Mau SL, Anderson MA, Cornish E, Tregear GW,
PI Crawford RJ, Niall HD, Bernatzky R;
DR WPI: 89-34992/48.
DR P-PSDB: P93469.
PT S-gene of a gametophytic self-incompatible plant - used to
PT produce S-protein and for genetic manipulation to create
PT self-incompatible cultivars
PS Example: Table 4, Page 20-21: 26pp: English.
CC The difference at the 3' end between clones NA-2-2 and NA-2-1
CC (see N92416) is near full-length but during sub-cloning a short
CC EcoRI fragment at the 5' end of the clone was inadvertently deleted.
CC The inventors claim a method for it isolating and identifying a cDNA
CC clone of an S-gene of a gametophytic self-incompatible plant (GSIP).
SQ Sequence 794 BP; 309 A; 136 C; 140 G; 207 T;
```

```
Query Match 1.2%; Score 36; DB 1; Length 794;
Best Local Similarity 100.0%; Pred. No. 6; 6e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 738 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 773

RESULT 31
X04314
ID X04314 standard; DNA: 808 BP.
AC X04314;
DT 13-APR-1999 (first entry)
DE Human secreted protein gene 4 clone HOVA158.
KM Human; secreted protein; fusion protein; gene therapy; protein therapy;
KM diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KM developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KM inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KM cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN M09856804-A1.
PD 17-DEC-1998.
PR 11-JUN-1998: U12125.
PR 02-OCT-1997: US-061060.
PR 13-JUN-1997: US-049547.
PR 13-JUN-1997: US-049548.
PR 13-JUN-1997: US-049549.
PR 13-JUN-1997: US-049550.
PR 13-JUN-1997: US-049606.
PR 13-JUN-1997: US-049607.
PR 13-JUN-1997: US-049608.
PR 13-JUN-1997: US-049609.
PR 13-JUN-1997: US-049610.
PR 13-JUN-1997: US-049611.
PR 13-JUN-1997: US-050566.
PR 13-JUN-1997: US-050901.
PR 13-JUN-1997: US-052989.
PR 08-JUL-1997: US-051919.
PR 18-AUG-1997: US-055984.
PR 12-SEP-1997: US-058665.
PR 12-SEP-1997: US-058666.
PR 12-SEP-1997: US-058667.
PR 12-SEP-1997: US-058750.
PR 12-SEP-1997: US-058972.
PR 12-SEP-1997: US-058972.
PR 12-SEP-1997: US-058975.
PR 02-OCT-1997: US-060834.
PR 02-OCT-1997: US-060841.
PR 02-OCT-1997: US-060844.
PR 02-OCT-1997: US-060865.
PR 02-OCT-1997: US-061059.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Ferrile AM, Feng P, Greene JM, Lafleur DM,
PI Moore PA, Nl J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P,
PI Yu GL.
DR WPI: 99-080881/07.
DR P-PSDB: W78129.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1: Page 169-170: 380pp: English.
CC This sequence represents a nucleic acid molecule which encodes a secreted
CC human protein. The gene number, and the clone it is derived from, are
CC detailed in the descriptor line. The gene can be used to generate fusion
CC proteins by linking to the gene to a human immunoglobulin Fc portion
CC (e.g. X04302) for increasing the stability of the fused protein as
CC compared to the human protein only.
CC The invention relates to 86 novel genes and their fragments (nucleic acid
CC sequences: X04311-X04410; amino acid sequences W78126-W78225) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 86 polynucleotides, based on
CC which tissues they are most highly expressed in (see X04311 for described
CC uses).
```

SQ Sequence 808 BP; 268 A; 165 C; 148 G; 225 T;

Query Match 1.2%; Score 36; DB 1; Length 808;

Best Local Similarity 100.0%; Pred. No. 6, 5e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

Db 756 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 791

RESULT 32

X51754

ID X51754 standard; DNA; 886 BP.

AC X51754;

DT 17-JUN-1999 (first entry)

DE DNA encoding a human secreted protein.

KW Human secreted protein; cancer; immune disorder; infection;

KW inflammatory disorder; skin disorder; tumour; atherosclerosis;

KW restenosis; autoimmune disorder; Alzheimer's disease;

KW peripheral neuropathy; trauma; spinal cord injury; allergy;

KW hematopoietic disorder; skeletal disorder; neurological disorder;

KW arthritic disorder; asthma; immunodeficiency disease; AIDS;

KW transplant rejection; ss.

OS Homo sapiens.

PN MO9911293-A1.

PD 11-MAR-1999.

PF 03-SEP-1998; U18360.

PR 12-SEP-1997; US-058974.

PR 05-SEP-1997; US-057626.

PR 05-SEP-1997; US-057663.

PR 05-SEP-1997; US-057669.

PR 12-SEP-1997; US-058666.

PR 12-SEP-1997; US-058667.

PR 12-SEP-1997; US-058973.

PA (HUMA-) HUMAN GENE SCIT INC.

PI Brewer LA, Ebner R, Lafleur DW, Moore PA, Olsen HS,

PI Rosen GA, Ruben SM, Shi Y;

DR WPI; 99-204988/17.

PT P-PSDB: Y12967.

PT New isolated human genes and the secreted polypeptides they encode

PT - useful for diagnosis and treatment of e.g. neurological disorders,

PT tumours, immune disorders, inflammation or haematological disorders,

PS Claim 1; Page 180; 215pp; English.

CC X51701-55 encode human secreted proteins. The polynucleotides and

CC their corresponding secreted polypeptides are useful for preventing,

CC treating or ameliorating medical conditions, e.g. by protein or gene

CC therapy. Pathological conditions can also be diagnosed by determining

CC the amount of the new polypeptides in a sample or by determining the

CC presence of mutations in the new polynucleotides. Specific uses are

CC described for each polynucleotide, based on which tissues they are

CC most highly expressed in, and include developing products for the

CC diagnosis or treatment of cancer, immune disorders, infection,

CC inflammatory disorders, skin disorders, tumours, atherosclerosis,

CC restenosis, autoimmune disorders, Alzheimer's disease, peripheral

CC neuropathies, trauma, spinal cord injuries, allergy, hematopoietic

CC disorders, skeletal disorders, neurological disorders, arthritic

CC disorders, asthma, immunodeficiency diseases, AIDS and transplant

CC rejection. The polypeptides are also useful for identifying their

CC binding partners.

SQ Sequence 886 BP; 288 A; 163 C; 174 G; 261 T;

Query Match 1.2%; Score 36; DB 1; Length 886;

Best Local Similarity 100.0%; Pred. No. 6, 5e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

Db 801 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 836

RESULT 33

ID Q92529 standard; cDNA; 893 BP.

AC Q92529;

DT 09-DEC-1995 (first entry)

DE P. communis (pear) arabinogalactan protein from cDNA clone PCAGP9.

KW Arabinogalactan protein; AGP; gum; thickening agent; emulsifier;

KW adhesive; ss.

OS Pyrus communis.

FT Key Location/Qualifiers

FT cds 70..507

FT /*tag= a

PN W09515377-A.

PD 08-JUN-1995.

PF 01-DEC-1994; AU0744.

PR 03-DEC-1993; US-161944.

PR 18-JUL-1994; US-276452.

PA (ALBR) ALBRIGHT & WILSON AUSTRALIA.

PI Batic A, Chen C, Clarke AE, Du H, Gane AM, Mau S;

PI WPI; 95-215258/28.

DR P-PSDB: R75511.

PT Recombinant plant arabinogalactan protein and cloned DNA encoding it

PT useful as an emulsifying agent, adhesive agent or a lubricating agent.

PS Claim 22; Figure 5A, 142pp; English.

CC Total native arabinogalactan protein (AGP) was isolated from

CC P. communis. It was deglycosylated using anhydrous HF.

CC The total deglycosylated AGP sample was then reduced and

CC carboxymethylated. It was then separated by HPLC. The fractions

CC were subjected to thermolysin digestion and then separated.

CC Individual peaks from thermolysis digested were sequenced.

CC The sequence R75536 was selected as a template for the isolation

CC or the corresp. AGP gene. An antisense RNA probe was synthesised

CC from the PCR template and used to screen a cDNA library prepd.

CC from pear cell suspension culture. Three cDNA clones were isolated

CC and sequenced. The sequence of the longest cDNA clone PCAGP9 is

CC shown in Q92529/R75511.

SQ Sequence 893 BP; 208 A; 248 C; 162 G; 275 T;

Query Match 1.2%; Score 36; DB 1; Length 893;

Best Local Similarity 100.0%; Pred. No. 6, 5e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

Db 855 TCAAAAAAAAAAAAAAAAAAAAAAAAAA 890

RESULT 34

ID V99723 standard; cDNA; 894 BP.

AC V99723;

DT 26-APR-1999 (first entry)

DE Human adult retina secreted protein bx200_13 cDNA.

KW Secreted protein; human; retina; bx200_13; ds.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 245..523

FT /*tag= a

PN W09856909-A2.

PD 17-DEC-1998.

PF 08-JUN-1998; U11822.

PR 05-JUN-1998; US-092722.

PR 11-JUN-1997; US-873218.

PA (GENM) GENETICS INST INC.

PI Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavallie ER,

PI McCoy JM, Racie LA, Spaulding V, Treacy M;

PI WPI; 99-080899/07.

DR P-PSDB: W95346.

PT New polynucleotides encoding secreted human proteins - derived from

PT human foetal brain, adult testes, foetal kidney, adult thyroid or

PT adult retina cDNA libraries

PS Claim 15; Page 74; 113pp; English.

CC This is the nucleotide sequence of cDNA clone bk200_13, which
 CC includes an open reading frame for a 92-amino acid polypeptide
 CC (see W95346). The clone was isolated from a human adult retina
 CC cDNA library using methods which are selective for cDNAs encoding
 CC secreted proteins, or was identified as encoding a secreted or
 CC transmembrane protein on the basis of computer analysis of the
 CC amino acid sequence of the encoding protein. Database searches
 CC indicate some sequence similarity to known sequences. The
 CC invention provides cDNA clones (see V99721-33) from human adult
 CC thymoid, adult retina, adult testis, foetal kidney and foetal brain
 CC that encode novel secreted proteins (see W95344-53). Each clone is
 CC individually available from deposit clone ATCC 98451 (see also
 CC V99734-43). The isolated polynucleotides (PNS) and proteins are
 CC predicted to have activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans
 CC and animals, although no supporting data is given. Suggested
 CC activities include nutritional, cytokine, cell proliferation or
 CC differentiation, immune stimulating (e.g. as vaccines) or immune
 CC suppressing, haematopoiesis regulating, tissue growth,
 CC activin/inhibin, chemotactic/chemokinetic, haemostatic,
 CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour
 CC invasion suppressor, and tumour inhibition activities. The PNS are
 CC also stated to be useful for gene therapy. 184 G; 319 T;
 SQ Sequence 894 BP; 274 A; 117 C; 184 G; 319 T;

Query Match 1.2%; Score 36; DB 1; Length 894;
 Best Local Similarity 100.0%; Pred. No. 6.5e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaa 3117
 Db 852 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 887

RESULT 35

ID X27330 standard; DNA; 921 BP.
 AC X27330;
 DT 11-JUN-1999 (first entry)
 DE Human secreted protein gene 20 clone HKAH10.
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 PN W09902546-A1.
 PD 21-JAN-1999.
 PF 07-JUL-1998; U13684.
 PR 12-SEP-1997; US-058785.
 PR 08-JUL-1997; US-051916.
 PR 08-JUL-1997; US-051918.
 PR 08-JUL-1997; US-051919.
 PR 08-JUL-1997; US-051920.
 PR 08-JUL-1997; US-051925.
 PR 08-JUL-1997; US-051926.
 PR 08-JUL-1997; US-051928.
 PR 08-JUL-1997; US-051928.
 PR 08-JUL-1997; US-051929.
 PR 08-JUL-1997; US-051930.
 PR 08-JUL-1997; US-051931.
 PR 08-JUL-1997; US-051932.
 PR 08-JUL-1997; US-052732.
 PR 08-JUL-1997; US-052733.
 PR 08-JUL-1997; US-052733.
 PR 08-JUL-1997; US-052793.
 PR 08-JUL-1997; US-052795.
 PR 08-JUL-1997; US-052803.
 PR 18-AUG-1997; US-055684.
 PR 18-AUG-1997; US-055722.
 PR 18-AUG-1997; US-055723.

PR 18-AUG-1997; US-055947.
 PR 18-AUG-1997; US-055948.
 PR 18-AUG-1997; US-055949.
 PR 18-AUG-1997; US-055950.
 PR 18-AUG-1997; US-055953.
 PR 18-AUG-1997; US-055954.
 PR 18-AUG-1997; US-055964.
 PR 18-AUG-1997; US-055984.
 PR 18-AUG-1997; US-056360.
 PR 12-SEP-1997; US-058660.
 PR 12-SEP-1997; US-058661.
 PR 12-SEP-1997; US-058664.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Olsen HS, Rosen CA, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA,
 PI Brewer LA, Ebner R, Ruben SM, Shi Y, Soppet DR, Zeng Z;
 DR WPI; 99-120770/10.
 DR P-PSDB; Y02669.

PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 PS Claim 1: Page 256: 464pp; English.
 CC This sequence represents a nucleic acid molecule which encodes a secreted
 CC human protein. The gene number, and the clone it is derived from, are
 CC detailed in the descriptor line. The gene can be used to generate fusion
 CC proteins by linking to the gene to a human immunoglobulin Fc portion
 CC (e.g. X27302) for increasing the stability of the fused protein as
 CC compared to the human protein only.
 CC The invention relates to 123 novel genes and their fragments (nucleic
 CC acid sequences: X27311-X27449; amino acid sequences Y02650-Y02788) which
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 123 polynucleotides, based on
 CC which tissues they are most highly expressed in (see X27311 for described
 CC uses).
 SQ Sequence 921 BP; 258 A; 187 C; 195 G; 281 T;

Query Match 1.2%; Score 36; DB 1; Length 921;
 Best Local Similarity 100.0%; Pred. No. 6.5e-05;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaa 3117
 Db 875 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 910

RESULT 36

ID X51728 standard; DNA; 921 BP.
 AC X51728;
 DT 17-JUN-1999 (first entry)
 DE DNA encoding a human secreted protein.
 KW Human secreted protein; cancer; immune disorder; infection;
 KW inflammatory disorder; skin disorder; tumour; atherosclerosis;
 KW restenosis; autoimmune disorder; Alzheimer's disease;
 KW peripheral neuropathy; trauma; spinal cord injury; allergy;
 KW hematopoietic disorder; skeletal disorder; neurological disorder;
 KW arthritic disorder; asthma; immunodeficiency disease; AIDS;
 KW transplant rejection; ss.
 OS Homo sapiens.
 PN W09911293-A1.
 PD 11-MAR-1999.
 PF 03-SEP-1998; U18360.
 PR 12-SEP-1997; US-058974.
 PR 05-SEP-1997; US-057626.
 PR 05-SEP-1997; US-057663.
 PR 05-SEP-1997; US-057669.
 PR 12-SEP-1997; US-058666.
 PR 12-SEP-1997; US-058667.
 PR 12-SEP-1997; US-058973.
 PA (HUMA-) HUMAN GENOME SCI INC.


```
PR 12-SEP-1997; US-058974.
PR 05-SEP-1997; US-057626.
PR 05-SEP-1997; US-057663.
PR 05-SEP-1997; US-057669.
PR 12-SEP-1997; US-058666.
PR 12-SEP-1997; US-058667.
PR 12-SEP-1997; US-058973.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Lafleur DW, Moore PA, Olsen HS,
PI Rosen CA, Ruben SM, Shi Y;
DR WPI: 99-204988/17.
DR P-PSDB: Y12961.
PT New isolated human genes and the secreted polypeptides they encode
PT - useful for diagnosis and treatment of e.g. neurological disorders,
PT tumours, immune disorders, inflammation or haematological disorders,
PS Claim 1; Page 177; 215pp; English.
CC X51701-55 encode human secreted proteins. The polynucleotides and
CC their corresponding secreted polypeptides are useful for preventing,
CC treating or ameliorating medical conditions, e.g. by protein or gene
CC therapy. Pathological conditions can also be diagnosed by determining
CC the amount of the new polypeptides in a sample or by determining the
CC presence of mutations in the new polynucleotides. Specific uses are
CC described for each polynucleotide, based on which tissues they are
CC most highly expressed in, and include developing products for the
CC diagnosis or treatment of cancer, immune disorders, infection,
CC inflammatory disorders, skin disorders, tumours, atherosclerosis,
CC restenosis, autoimmune disorders, Alzheimer's disease, peripheral
CC neuropathies, trauma, spinal cord injuries, allergy, hematopoietic
CC disorders, skeletal disorders, neurological disorders, arthritic
CC disorders, asthma, immunodeficiency diseases, AIDS and transplant
CC rejection. The polypeptides are also useful for identifying their
CC binding partners.
SQ Sequence 990 BP; 320 A; 169 C; 159 G; 335 T;

Query Match 1.2%; Score 36; DB 1; Length 990;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 929 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 964

RESULT 40
ID X37470 standard; cDNA; 1008 BP.
AC X37470;
DE 06-JUL-1999 (first entry)
KW Human secreted protein cDNA fragment containing gene 20.
KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW immune system disease; autoimmune disease; hepatic disease; lymphoma;
KW renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;
KW cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;
KW pulmonary disorder; transplant rejection; osteoclast; osteoporosis;
KW arthritis; malignancy; digestive; endocrine; infection; ss.
OS Homo sapiens.
PN W09918208-Al.
PD 15-APR-1999.
PF 01-OCT-1998; U20775.
PR 02-OCT-1997; US-060884.
PR 02-OCT-1997; US-060833.
PR 02-OCT-1997; US-060836.
PR 02-OCT-1997; US-060837.
PR 02-OCT-1997; US-060838.
PR 02-OCT-1997; US-060839.
PR 02-OCT-1997; US-060843.
PR 02-OCT-1997; US-060862.
PR 02-OCT-1997; US-060866.
PR 02-OCT-1997; US-060874.
PA (HUMA-) HUMAN GENOME SCI INC.
```

```
PI Carter KC, Duan DR, Endress GA, Feng P, Ferrie AM,
PI Florence KA, Greene JM, Janat F, Lafleur DW, Ni J,
PI Rosen CA, Ruben SM, Shi Y, Young P, Yu G;
DR WPI: 99-264022/22.
DR P-PSDB: Y07871.
PT New isolated human genes and the secreted polypeptides they encode
PS Claim 1a; Page 238-239; 368pp; English.
CC This invention describes novel isolated human genes and the secreted
CC proteins they encode. The products of the invention are useful for
CC preventing, treating or ameliorating medical conditions, e.g. by protein
CC or gene therapy. Also pathological conditions can be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 101 polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours,
CC neurodegenerative disorders, developmental abnormalities and fetal
CC deficiencies, blood disorders, leukemias, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,
CC allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate
CC disease, skeletal or cardiac muscle disorders, pulmonary disorders,
CC transplant rejection, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, digestive/endocrine disorders,
CC infections and AIDS. The human secreted proteins of the invention are
CC represented in Y07852-Y07993 and the encoding nucleic acids are
CC represented in X37451-X37552.
SQ Sequence 1008 BP; 273 A; 204 C; 190 G; 340 T;

Query Match 1.2%; Score 36; DB 1; Length 1008;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 967 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1002

RESULT 41
ID V45444 standard; cDNA; 1029 BP.
AC V45444;
DE 02-FEB-1999 (first entry)
DE Human chemokine ZSIG-35 cDNA.
KW ZSIG-35; beta-chemokine; human; ligand; lymphocyte migration;
KW inflammation; ischaemia; reperfusion injury; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 102..554
FT sig_peptide 102..161
FT mat_peptide 162..551
FT WO9844117-Al.
PD 08-OCT-1998.
PF 27-MAR-1998; U06115.
PR 09-MAY-1997; US-046083.
PR 28-MAR-1997; US-042862.
PA (ZYMO ) ZYMOGENETICS INC.
PI Sheppard PO;
DR WPI: 98-557114/47.
DR P-PSDB: W30565.
PT New human chemokine ZSIG-35 - used for, e.g. treating inflammatory
PT disease, lymphocyte migration and ischaemia/reperfusion injury
PS Claim 15; Page 80-81; 105pp; English.
CC This nucleotide sequence includes an open reading frame for a novel
CC human chemokine (see W30565), designated ZSIG-35, that has homology
CC to members of the beta-chemokine family, in particular to murine
CC macrophage inflammatory protein. ZSIG-35 polynucleotides were
CC initially identified by querying an expressed sequence tag
CC database. The gene was mapped to chromosome 19p13.3. The
CC invention provides ZSIG-35 polynucleotides and polypeptides, an
```

CC expression vector, a cultured cell carrying the vector, and an
 CC oligonucleotide probe or primer (see V45446). The cultured cell
 CC may be used for the production of protein (Claimed).
 CC ZS1G-35 may be used in therapeutics (claimed) for the regulation of
 CC acute and chronic inflammatory disease conditions, lymphocyte
 CC migration and ischaemia/reperfusion injury. The probe can be used
 CC to detect a genetic abnormality in a patient (claimed) by comparing
 CC the hybridisation of a genetic sample to the probe to produce a
 CC first reaction product, and then comparing this product to a
 CC control.
 CC Sequence 1029 BP; 241 A; 303 C; 261 G; 224 T;

Query Match 1.2%; Score 36; DB 1; Length 1029;

Best Local Similarity 100.0%; Pred. No. 6.4e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaa 3117
 Db 977 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1012

RESULT 42

ID X30150 standard; DNA; 1037 BP.

AC X30150;

DT 18-JUN-1999 (first entry)

DE Human secreted protein gene 6.

KW Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; ss.

OS Homo sapiens.

PN W09910363-A1.

PD 04-MAR-1999.

PF 27-AUG-1998; U17709.

PR 29-AUG-1997; US-056271.

PR 29-AUG-1997; US-056073.

PR 29-AUG-1997; US-056247.

PR 29-AUG-1997; US-056270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Fan P, Kyaw H, Rosen CA, Ruben SM, Wei YF;

DR WPI: 99-190585/16.

DR F-PSDB; Y04298.

PT New isolated human genes and the secreted polypeptides they encode
 PT - useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders

PS Claim 1; Page 135; 170pp; English.

CC X30145 to X30173 represent 29 isolated human secreted protein genes.

CC Y04293 to Y04321 represent the secreted proteins encoded by the 29 human
 CC genes. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new genes. Specific
 CC uses are described for each of the 29 genes, based on which tissues they
 CC are most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
 CC disorders, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.

CC The sequences given in X30174 to X30182 and Y04322 to Y04334 are used

CC in the exemplification of the present invention.

CC Sequence 1037 BP; 308 A; 198 C; 224 G; 306 T;

Query Match

Best Local Similarity 100.0%; Score 36; DB 1; Length 1037;

Pred. No. 6.4e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaa 3117

Db 995 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1030

RESULT 43

ID Q10377 standard; DNA; 1046 BP.

AC Q10377;

DT 12-APR-1991 (first entry)

DE Plasmid pMG3C9 used to isolate style-stigma specific gene STG3C9.

KW Hybrid vigour; pollination; ss.

OS Nictiana tabacum "Petit Havana" SRL.

FH Key Location/Qualifiers

FT cds 3..803

FT /*tag= a

PN EP-412006-A.

PD 06-FEB-1991.

PF 31-JUL-1990; 402196.

PR 04-AUG-1989; EP-402224.

PR 31-JUL-1990; EP-402196.

PA (PLAN-) PLANT GENETIC SYST.

PI De Greef W, Van Emmelo J, De Oliveira DE, De Souza MH;

PI Van Montagu M;

DR WPI: 91-038878/06.

DR P-PSDB; R10535.

PT Modified plant with transformed nuclear genome - obt'd. by using
 PT foreign DNA encoding prod. which disrupts metabolism, functioning
 PT and/or development of plant.

PS Disclosure: Fig 2B; 42pp; English.

CC The clone was produced by subcloning a style-stigma specific cDNA,
 CC pMG3C9. The clone was found to be stigma-specific in
 CC Northern analysis. A probe from the clone was used to isolate

CC the corresp. genomic sequence which is specifically expressed in
 CC style-stigma tissues of the female organ of tobacco. The corresp.
 CC clone, pSTG3C9, contains STMG-type" gene. This and

CC other similar genes can be used to produce new hybrid plants or

CC seeds having a combination of desirable traits, and showing hybrid

CC vigour. A female-sterile, male-fertile plant can be produced which

CC favours cross-pollination.

CC See also Q10374-76.

CC Sequence 1046 BP; 360 A; 254 C; 166 G; 266 T;

Query Match

Best Local Similarity 1.2%; Score 36; DB 1; Length 1046;

Pred. No. 6.4e-05;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaa 3117

Db 964 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 999

RESULT 44

ID X37464

AC X37464 standard; cDNA; 1060 BP.

DT 06-JUL-1999 (first entry)

DE Human secreted protein cDNA fragment containing gene 14.

KW Human; secreted protein; treatment; prevention; protein therapy; AIDS;

KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;

KW developmental abnormality; fetal deficiency; blood disorder; leukemia;

KW immune system disease; autoimmune disease; hepatic disease; lymphoma;

KW renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;

KW cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;

KW pulmonary disorder; transplant rejection; osteoclast; osteoporosis;

KW arthritis; malignancy; digestive; endocrine; infection; ss.

OS Homo sapiens.

PN W09918208-A1.

PD 15-APR-1999.


```
PF 01-OCT-1998; U20775.
PR 02-OCT-1997; US-060884.
PR 02-OCT-1997; US-060833.
PR 02-OCT-1997; US-060836.
PR 02-OCT-1997; US-060837.
PR 02-OCT-1997; US-060838.
PR 02-OCT-1997; US-060839.
PR 02-OCT-1997; US-060843.
PR 02-OCT-1997; US-060862.
PR 02-OCT-1997; US-060866.
PR 02-OCT-1997; US-060874.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Carter KC, Duan DR, Endress GA, Feng P, Ferrie AM,
PI Florence KA, Greene JM, Janat F, Lafleur DW, Ni J,
PI Rosen CA, Ruben SM, Shi Y, Young P, Yu G;
DR WPI: 99-264022/22.
DR P-PSDB; Y07865.
DR P-PSDB; Y07865.
PT New isolated human genes and the secreted polypeptides they encode
PS Claim 1a; Page 235-236; 369pp; English.
CC This invention describes novel isolated human genes and the secreted
CC proteins they encode. The products of the invention are useful for
CC preventing, treating or ameliorating medical conditions, e.g. by protein
CC or gene therapy. Also pathological conditions can be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 101 polynucleotides,
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours,
CC neurodegenerative disorders, developmental abnormalities and fetal
CC deficiencies, blood disorders, leukemias, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,
CC allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate
CC disease, skeletal or cardiac muscle disorders, pulmonary disorders,
CC transplant rejection, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, digestive/endocrine disorders,
CC infections and AIDS. The human secreted proteins of the invention are
CC represented in Y07852-Y07993 and the encoding nucleic acids are
CC represented in X37451-X37552.
SQ Sequence 1060 BP; 328 A; 205 C; 192 G; 326 T;

Query Match 1.2%; Score 36; DB 1; Length 1060;
Best Local Similarity 100.0%; Pred. No. 6.3e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 958 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 993

RESULT 45
X37460
ID X37460 standard; cDNA; 1089 BP.
AC X37460;
DT 06-JUL-1999 (first entry)
DE Human secreted protein cDNA fragment containing gene 10.
KW Human; secreted protein; treatment; prevention; protein therapy; AIDS;
KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW immune system disease; autoimmune disease; hepatic disease; lymphoma;
KW renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;
KW cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;
KW pulmonary disorder; transplant rejection; osteoclast; osteoporosis;
KW arthritis; malignancy; digestive; endocrine; infection; ss.
OS Homo sapiens.
PN WO9918208-A1.
PD 15-APR-1999.
PF 01-OCT-1998; U20775.
PR 02-OCT-1997; US-060884.
PR 02-OCT-1997; US-060833.
PR 02-OCT-1997; US-060836.
PR 02-OCT-1997; US-060837.
PR 02-OCT-1997; US-060838.

101-OCT-1998; U20775.
PR 02-OCT-1997; US-060839.
PR 02-OCT-1997; US-060843.
PR 02-OCT-1997; US-060862.
PR 02-OCT-1997; US-060866.
PR 02-OCT-1997; US-060874.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Carter KC, Duan DR, Endress GA, Feng P, Ferrie AM,
PI Florence KA, Greene JM, Janat F, Lafleur DW, Ni J,
PI Rosen CA, Ruben SM, Shi Y, Young P, Yu G;
DR WPI: 99-264022/22.
DR P-PSDB; Y07861.
DR P-PSDB; Y07861.
PT New isolated human genes and the secreted polypeptides they encode
PS Claim 1a; Page 232-233; 368pp; English.
CC This invention describes novel isolated human genes and the secreted
CC proteins they encode. The products of the invention are useful for
CC preventing, treating or ameliorating medical conditions, e.g. by protein
CC or gene therapy. Also pathological conditions can be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 101 polynucleotides,
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours,
CC neurodegenerative disorders, developmental abnormalities and fetal
CC deficiencies, blood disorders, leukemias, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,
CC allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate
CC disease, skeletal or cardiac muscle disorders, pulmonary disorders,
CC transplant rejection, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, digestive/endocrine disorders,
CC infections and AIDS. The human secreted proteins of the invention are
CC represented in Y07852-Y07993 and the encoding nucleic acids are
CC represented in X37451-X37552.
SQ Sequence 1089 BP; 367 A; 202 C; 240 G; 279 T;

Query Match 1.2%; Score 36; DB 1; Length 1089;
Best Local Similarity 100.0%; Pred. No. 6.3e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3082 tcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
DB 1042 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1077

Search completed: September 11, 2000, 00:14:14
Job time: 7147 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2000, 21:40:42 ; Search time 1087.01 Seconds
(without alignments)
12645.178 Million cell updates/sec

Title: US-09-454-334-1
Perfect score: 3117
Sequence: 1 ccgggtgaccacgcgctcg.....aaaaaaaaaaaaaaaaaaaaa 3117

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5247842 seqs, 2204914090 residues

Word size : 0
Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Listing first 1000 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_gss1:*
94: gb_gss2:*
95: gb_gss3:*
96: gb_gss4:*
97: em_gss1:*
98: em_gss2:*
99: em_gss3:*
100: em_gss4:*
101: gb_gss5:*
102: gb_gss6:*
103: gb_gss7:*
104: gb_gss8:*
105: gb_gss9:*
106: em_gss5:*
107: em_gss6:*
108: em_gss7:*
109: em_gss8:*
110: em_gss9:*
111: em_gss10:*
112: em_gss11:*
113: gb_gss10:*
114: gb_gss11:*
115: em_gss12:*
116: gb_gss12:*

117: gb_gss13:*
118: gb_gss14:*
119: gb_gss15:*
120: gb_gss16:*
121: gb_gss17:*
122: gb_gss18:*
123: gb_gss19:*
124: em_gss13:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	576	18.5	657	AI013917	AI013917 UT208592
C 2	65	2.1	387	AA925917	AA925917 UT-R-A1-e
C 3	61	2.0	496	AI891958	AI891958 ul57a09.y
C 4	52	1.7	349	AW325627	AW325627 17814 MAR
C 5	45	1.4	442	AA980401	AA980401 ua52911.r
C 6	41	1.3	157	AL040100	AL040100 DKF2p434B
C 7	41	1.3	157	AW278153	AW278153 sf40808.y
C 8	41	1.3	240	AU073712	AU073712 AU073712
C 9	41	1.3	262	AU065863	AU065863 AU065863
C 10	41	1.3	1367	AF034791	AF034791 AF034791
C 11	40	1.3	210	AW313563	AW313563 8840 MARC
C 12	40	1.3	224	AW657118	AW657118 109764 MA
C 13	40	1.3	275	AI896601	AI896601 C98601 Rice
C 14	40	1.3	449	AI253246	AI253246 qz39e12.x
C 15	39	1.3	79	AI320733	AI320733 C9f03nm.f
C 16	39	1.3	139	AI946912	AI946912 ps32h08.y
C 17	39	1.3	161	AI540624	AI540624 tn88609.x
C 18	39	1.3	167	AA231085	AA231085 mw11e11.r
C 19	39	1.3	182	AI630474	AI630474 ad11a06.y
C 20	39	1.3	190	AW477617	AW477617 15122 MAR
C 21	39	1.3	207	AW172002	AW172002 618047C10
C 22	39	1.3	229	AI357288	AI357288 40265 MAR
C 23	39	1.3	245	AA832425	AA832425 cc99d12.s
C 24	39	1.3	281	AW056177	AW056177 660005C01
C 25	39	1.3	307	AW479416	AW479416 25195 MAR
C 26	39	1.3	318	AA237390	AA237390 mx18b02.r
C 27	39	1.3	327	AU082521	AU082521 AU082521
C 28	39	1.3	337	AA773444	AA773444 ad58d05.s
C 29	39	1.3	340	AA595839	AA595839 nn05c01.s
C 30	39	1.3	341	AW394641	AW394641 sh06a07.y
C 31	39	1.3	376	AI529904	AI529904 ui83h01.y
C 32	39	1.3	380	D12652	D12652 RICC0739B R
C 33	39	1.3	424	AI829162	AI829162 wk76a05.x
C 34	39	1.3	427	AI050855	AI050855 oy47a05.x
C 35	39	1.3	444	AA610049	AA610049 af18h01.s
C 36	39	1.3	450	AA731682	AA731682 nw58e10.s
C 37	39	1.3	475	AW249071	AW249071 2820891.5
C 38	39	1.3	482	AI018664	AI018664 ov65e05.s
C 39	39	1.3	499	AA393950	AA393950 zt78a10.r
C 40	39	1.3	506	AI31	AA808304 oc41f12.s
C 41	39	1.3	510	AA810286	AA810286 oc41h11.s
C 42	39	1.3	519	AI241203	AI241203 qj96g10.x
C 43	39	1.3	520	AT000142	AT000142 AT000142
C 44	39	1.3	523	AI397624	AI397624 NCSC5611T
C 45	39	1.3	523	AA566146	AA566146 660002F05
C 46	39	1.3	738	AI187484	AI187484 EST273 Ma
C 47	38	1.2	96	AA607306	AA607306 vm96f10.r
C 48	38	1.2	115	AI619093	AI619093 486073C06
C 49	38	1.2	145	AA829401	AA829401 q060406.s
C 50	38	1.2	166	AI567226	AI567226 tp51f02.x
C 51	38	1.2	169	AW164849	AW164849 se78a05.y
C 52	38	1.2	179	AA477009	AA477009 g9a0604.y
C 53	38	1.2	188	AA280902	AA280902 z897f10.s
C 54	38	1.2	188	AA497258	AA497258 fa04e08.s
C 55	38	1.2	189	AA656645	AA656645 109003 MA

38	1.2	193	AW100662	AW100662
38	1.2	203	AA485903	AA485903
38	1.2	208	AA869427	AA869427 vc07h10.r
38	1.2	234	AW583947	AW583947 ia07d01.x
38	1.2	239	AW102081	AW102081 SF282908.y
38	1.2	242	AW278164	AW278164 sf40e07.y
38	1.2	244	AI695228	AI695228 wa01g09.x
38	1.2	245	AA738321	AA738321 nx16h11.s
38	1.2	246	AI092654	AI092654 qa80b07.x
38	1.2	264	AU082927	AU082927
38	1.2	272	AA711214	AA711214 vt70b06.r
38	1.2	278	AI491852	AI491852 to01e04.x
38	1.2	280	AA4930261	AA4930261 oc45f08.s
38	1.2	284	C97620	C97620 C97620 Rice
38	1.2	285	AI563599	AI563599 vx91b04.x
38	1.2	289	AU033170	AU033170
38	1.2	297	AU033074	AU033074
38	1.2	302	AU030838	AU030838
38	1.2	304	AW311758	AW311758 5631 MARC
38	1.2	318	AA154878	AA154878 mx32e05.r
38	1.2	322	AI808311	AI808311 oc41g07.s
38	1.2	322	AI166640	AI166640 xg54g08.x
38	1.2	323	AI499147	AI499147 to05g09.x
38	1.2	325	AW021091	AW021091 df18f09.y
38	1.2	329	W79740	W79740 zd86g04.r1
38	1.2	337	AI896639	AI896639 C98639 Rice
38	1.2	338	AW733654	AW733654 sr83c12.y
38	1.2	341	AW290909	AW290909 nw25g12.s
38	1.2	341	C93971	C93971 C93971 D1ct
38	1.2	343	AW075394	AW075394 AA075394
38	1.2	346	AW734302	AW734302 sk81d04.y
38	1.2	352	AI893307	AI893307 C99307 Rice
38	1.2	353	AW307000	AW307000 sf51e06.y
38	1.2	356	AA277361	AA277361 va78g10.r
38	1.2	356	AA0771018	AA0771018
38	1.2	357	AA837811	AA837811 oe39g07.s
38	1.2	363	AA871182	AA871182 vg32b04.r
38	1.2	369	AI681953	AI681953 tp45a07.x
38	1.2	372	AI822331	AI822331 LO-776T3
38	1.2	373	AA609483	AA609483 afl4b02.s
38	1.2	384	AA198948	AA198948 mw39g03.r
38	1.2	390	AA495424	AA495424 fa03b01.s
38	1.2	402	AW311095	AW311095 sg32g08.y
38	1.2	411	AA129588	AA129588 z11ld09.s
38	1.2	414	AA217736	AA217736 mw56g03.r
38	1.2	417	AA127454	AA127454 zn29f10.s
38	1.2	423	AA831663	AA831663 oc83g12.s
38	1.2	431	N26206	N26206 yw53a09.s1
38	1.2	435	AA281501	AA281501 zt08g04.s
38	1.2	448	AU029299	AU029299
38	1.2	450	AU039070	AU039070
38	1.2	467	AA131365	AA131365 zo08b06.s
38	1.2	483	AA810223	AA810223 od14a09.s
38	1.2	484	AI973520	AI973520 496027F11
38	1.2	486	AA836268	AA836268 od17e05.s
38	1.2	490	AA832000	AA832000 oc97c11.s
38	1.2	497	AA603342	AA603342 np29d07.s
38	1.2	510	AA827614	AA827614 of04h03.s
38	1.2	521	AA678835	AA678835 ah09d04.s
38	1.2	533	AA490301	AA490301 aa51a03.s
38	1.2	548	AA710038	AA710038 vt38g03.r
38	1.2	548	AI257985	AI257985 LP06506.s
38	1.2	548	AU058000	AU058000
38	1.2	548	AW186926	AW186926 BNLGHI738
38	1.2	552	AA825404	AA825404 oc57h10.s
38	1.2	563	AU056286	AU056286
38	1.2	583	AI398912	AI398912 NCW06b6T7
38	1.2	592	AU086008	AU086008
38	1.2	615	AA108889	AA108889 mx156b11.r
38	1.2	628	AW090702	AW090702 xc95b11.x
38	1.2	637	AA189241	AA189241 x103b12.x
38	1.2	653	AA837091	AA837091 od18h01.s
38	1.2	703	AU030679	AU030679

129	38	1.2	706	81	C97853	C97853 C97853 Rice	c	202	37	1.2	256	29	AA643982	AA643982 nj40h09.s
130	37	1.2	50	28	AA590944	AA590944 vm25f02.r	203	37	1.2	256	47	AU029273	AU029273 AU029273	
131	37	1.2	56	79	AW632959	AW632959 bl02c02.x	204	37	1.2	258	72	AW424013	AW424013 sh59q12.y	
132	37	1.2	69	28	AA607058	AA607058 vm95e07.r	205	37	1.2	262	63	AW099006	AW099006 sd33a03.y	
133	37	1.2	69	31	AA795179	AA795179 vq94g07.r	206	37	1.2	263	72	AW432647	AW432647 sh83e05.y	
134	37	1.2	77	89	R99209	R99209 yq64e12.r1	207	37	1.2	264	47	AU029317	AU029317 AU029317	
135	37	1.2	88	77	HS0002931	AI038455 Homo sapi	208	37	1.2	267	35	AI130479	SWOVL3CAN	
136	37	1.2	98	38	AI349279	AI349279 ta75d08.x	209	37	1.2	269	72	AW424164	AW424164 sh61q07.y	
137	37	1.2	100	36	AI210082	AI210082 g9c01a1.r	c	210	37	1.2	270	34	AI007951	EST202402
138	37	1.2	102	36	AI676803	AI676803 oa45h12.s	211	37	1.2	270	70	AW311267	sg34f02.y	
139	37	1.2	104	69	AW191003	AW191003 xl67a03.x	212	37	1.2	272	48	AU058132	AU058132 AU058132	
140	37	1.2	109	73	AW481860	AW481860 38894 MAR	213	37	1.2	273	69	AW185815	se60a06.y	
141	37	1.2	110	43	AI702343	AI702343 tz66b01.x	c	214	37	1.2	274	41	AI596861	vs32a05.x
142	37	1.2	112	38	AI340552	AI340552 tb30f07.x	215	37	1.2	275	37	AI251573	fw45a10.x	
143	37	1.2	112	38	AI340553	AI340553 tb30h09.x	c	216	37	1.2	275	41	AI591131	tw23f11.x
144	37	1.2	125	41	AI538850	AI538850 tp74e02.x	217	37	1.2	275	64	AW126272	N100379e	
145	37	1.2	127	21	AA098579	AA098579 mm83f05.r	218	37	1.2	275	72	AW459492	sh41h11.y	
146	37	1.2	128	38	AI335363	AI335363 tb79a05.x	219	37	1.2	276	48	AU085927	AU085927 AU085927	
147	37	1.2	131	38	AI343091	AI343091 tb04b01.x	220	37	1.2	277	30	AA734166	AW734166 vs19h02.r	
148	37	1.2	134	64	AW156654	AW156654 se29a07.y	c	221	37	1.2	278	64	AW129218	xf21e11.x
149	37	1.2	136	46	AI950675	AI950675 wx52h04.x	222	37	1.2	282	72	AW433419	AW433419 707068G03	
150	37	1.2	137	22	AA191932	AA191932 rs04b02.r	c	223	37	1.2	284	41	AI545574	fb70f08.x
151	37	1.2	138	37	AI273966	AI273966 qu90h10.x	224	37	1.2	284	47	AU033071	AU033071 AU033071	
152	37	1.2	145	42	AI649528	AI649528 603005G08	c	225	37	1.2	286	33	AA987387	q986g10.s
153	37	1.2	148	43	AI687016	AI687016 tp81h10.x	c	226	37	1.2	289	46	AI932547	AI932547 wo03f11.x
154	37	1.2	154	63	AW099488	AW099488 sd41d09.y	227	37	1.2	290	70	AW306179	AW306179 se47a01.y	
155	37	1.2	154	73	AW480921	AW480921 33577 MAR	228	37	1.2	290	81	C93414	C93414 C93414 D1ct	
156	37	1.2	155	22	AA162707	AA162707 mn42c02.r	c	229	37	1.2	291	31	AA807269	oc33d07.s
157	37	1.2	155	47	AI036886	AI036886 DKF2p564P	c	230	37	1.2	291	40	AI492950	qz47a07.x
158	37	1.2	157	64	AW158429	AW158429 se26a12.y	231	37	1.2	291	46	AI944478	bs01h05.y	
159	37	1.2	161	29	AA662912	AA662912 nu92b02.s	c	232	37	1.2	295	31	AA837029	og09e06.s
160	37	1.2	161	45	AI1886062	AI1886062 wn15a04.x	c	233	37	1.2	296	22	AA8185548	mu48h04.r
161	37	1.2	162	42	AI666647	AI666647 mt53b05.x	c	234	37	1.2	297	91	W72952	zd58a09.s1
162	37	1.2	169	32	AA873345	AA873345 ob64g08.s	c	235	37	1.2	299	43	AI739780	606051A01
163	37	1.2	170	31	AA827658	AA827658 od01c02.s	236	37	1.2	302	64	AW132999	se13b01.y	
164	37	1.2	171	26	AA414389	AA414389 vd07f10.s	237	37	1.2	302	79	AW640213	AW640213 bl92e11.w	
165	37	1.2	172	22	AA162226	AA162226 mn44g01.r	238	37	1.2	302	81	C90617	C90617 D1ct	
166	37	1.2	176	63	AW100657	AW100657 sd58c08.y	c	239	37	1.2	313	46	AI917106	AI917106 ts74d03.x
167	37	1.2	182	28	AA606861	AA606861 vm83b09.r	c	240	37	1.2	315	27	AA517991	vl18c06.r
168	37	1.2	183	89	TO1212	TO1212 wEST01933 E	c	241	37	1.2	315	46	AI926024	wo41a05.x
169	37	1.2	184	72	AW414544	AW414544 48060 MAR	242	37	1.2	322	48	AU075938	AU075938 AU075938	
170	37	1.2	189	72	AA332419	AA332419 sh73g02.y	c	243	37	1.2	322	69	AW193598	AW193598 xm18b01.x
171	37	1.2	193	71	AW396099	AW396099 sh02c11.y	244	37	1.2	323	30	AA709810	AA709810 vt39a06.r	
172	37	1.2	194	40	AI492521	AI492521 t129a02.x	245	37	1.2	326	72	AW458018	sh77b11.y	
173	37	1.2	194	63	AW101857	AW101857 sh71e12.y	c	246	37	1.2	327	31	AA775857	ad17f03.s
174	37	1.2	199	20	AA017992	AA017992 mh45a07.r	247	37	1.2	327	42	AI649540	AI649540 603006A06	
175	37	1.2	199	47	AJ012951	AJ012951 EHST101	248	37	1.2	329	20	AA024135	mn97a02.r	
176	37	1.2	201	73	AW503171	AW503171 UI-HF-BNO	249	37	1.2	330	91	WB4086	WB4086 T2987 MVA74	
177	37	1.2	204	31	AA836664	AA836664 of53a03.s	c	250	37	1.2	331	23	AA260469	va94d04.r
178	37	1.2	205	41	AI540361	AI540361 tp63h08.x	c	251	37	1.2	337	28	AA558973	nl10c02.s
179	37	1.2	207	50	AV173975	AV173975 AV173975	252	37	1.2	337	81	C93386	C93386 C93386 D1ct	
180	37	1.2	211	45	AI862245	AI862245 tb77c12.x	253	37	1.2	337	81	C93399	C93399 C93399 D1ct	
181	37	1.2	211	62	AW026635	AW026635 ww45e08.x	c	254	37	1.2	338	46	AI927899	wp93b12.x
182	37	1.2	212	38	AI345358	AI345358 tb68e12.x	c	255	37	1.2	338	71	AW396703	sh29f10.y
183	37	1.2	212	87	N98121	N98121 2213C3 czap	c	256	37	1.2	344	31	AA830625	AA830625 oc50g09.s
184	37	1.2	213	20	AA018016	AA018016 mh45c07.r	c	257	37	1.2	344	38	AI3331810	AI3331810 fb01q12.x
185	37	1.2	217	46	AI946027	AI946027 bs19g10.y	c	258	37	1.2	346	32	AA918425	ol70h02.s
186	37	1.2	221	27	AA523466	AA523466 ni63h06.s	c	259	37	1.2	346	63	AA767170	AA767170 oc88h08.s
187	37	1.2	225	64	AW132200	AW132200 sd67e05.y	c	260	37	1.2	353	30	AW733209	AW733209 sk71d05.y
188	37	1.2	233	63	AW080611	AW080611 xc44a06.x	261	37	1.2	356	80	AI086823	AI086823 oz57g01.x	
189	37	1.2	236	23	AA274698	AA274698 vc05b05.r	c	262	37	1.2	357	35	AW15996	AW15996 50650 MAR
190	37	1.2	238	73	AA290000	AA290000 vb50b07.r	263	37	1.2	357	72	AW415996	AW415996 50650 MAR	
191	37	1.2	239	73	AA483334	AA483334 52580 MAR	264	37	1.2	359	81	C91502	C91502 D1ct	
192	37	1.2	240	48	AU072476	AU072476 AU072476	c	265	37	1.2	361	92	X73770	X73770 CHESTW067 G
193	37	1.2	240	48	AU074000	AU074000 AU074000	266	37	1.2	362	47	AU071104	AU071104 AU071104	
194	37	1.2	244	42	AI650711	AI650711 wb25g07.x	267	37	1.2	363	47	AU029668	AU029668 AU029668	
195	37	1.2	245	63	AW100838	AW100838 sd62a06.y	c	268	37	1.2	363	70	AW288760	AW288760 707009D01
196	37	1.2	245	87	N69347	N69347 yz43e08.s1	269	37	1.2	364	35	AI073388	AI073388 oo13a08.x	
197	37	1.2	248	64	AW119367	AW119367 sd45c07.y	c	270	37	1.2	366	38	AI341886	AI341886 qs98b09.x
198	37	1.2	251	88	R65559	R65559 14063 Lambd	271	37	1.2	366	92	Z98088	Z98088 Z98088 Phys	
199	37	1.2	253	40	AI495706	AI495706 sb15c08.y	c	272	37	1.2	369	39	AI395143	AI395143 WAO02517
200	37	1.2	253	64	AW119354	AW119354 sd45a12.y	273	37	1.2	369	70	AW288761	AW288761 707009D01	
201	37	1.2	255	63	AW065230	AW065230 614038B03	274	37	1.2	369	81	C99552	C99552 C99552 Rice	

275	37	1.2	370	47	AU031086	AU031086	AU031086	C 348	37	1.2	501	63	AW058793	AW058793	fe47h06.x
276	37	1.2	371	30	AA760782	nz14c04.s	AA760782	C 349	37	1.2	502	69	AW180895	MgA0006r	
277	37	1.2	372	37	AI271716	qj87g06.x	AI271716	C 350	37	1.2	511	29	AA633412	np78e10.s	
278	37	1.2	375	47	AU029637	AU029637	AU029637	C 351	37	1.2	512	38	AI382230	td04305.x	
279	37	1.2	378	31	AA804629	ob98c12.s	AA804629	C 352	37	1.2	513	31	AA829330	od05h09.s	
280	37	1.2	378	64	AW129589	xel9e04.x	AW129589	C 353	37	1.2	513	31	AA830374	oc64g01.s	
281	37	1.2	381	23	AA259538	va48h11.r	AA259538	C 354	37	1.2	516	45	AI738199	606047A12	
282	37	1.2	383	30	AA709128	zf92h07.s	AA709128	C 355	37	1.2	518	45	AI890391	vm84g04.x	
283	37	1.2	383	39	AI395176	MA002571.	AI395176	C 356	37	1.2	521	30	AA766930	oc88f12.s	
284	37	1.2	384	48	AU057780	AU057780	AU057780	C 357	37	1.2	532	48	AU082640	AU082640	
285	37	1.2	385	23	AA243164	zr26d01.s	AA243164	C 358	37	1.2	533	48	AU082139	AU082139	
286	37	1.2	385	79	AW669208	112464 MA	AW669208	C 359	37	1.2	537	31	AA835963	oc79h02.s	
287	37	1.2	387	48	AU071229	AU071229	AU071229	C 360	37	1.2	538	21	AA081411	zn34h10.s	
288	37	1.2	389	39	AI393637	tg59e12.x	AI393637	C 361	37	1.2	539	42	AI666130	606005C10	
289	37	1.2	389	79	AW633117	bl04a02.x	AW633117	C 362	37	1.2	547	28	AA547940	nk51h07.s	
290	37	1.2	397	63	AW078751	xb33h03.x	AW078751	C 363	37	1.2	547	28	AA594974	nc37b09.s	
291	37	1.2	399	31	AA777002	zr23f09.s	AA777002	C 364	37	1.2	549	43	AI728861	BNLGH1118	
292	37	1.2	399	37	AI226246	ue88h04.y	AI226246	C 365	37	1.2	552	48	AU037695	AU037695	
293	37	1.2	399	74	AW568851	sl73a10.y	AW568851	C 366	37	1.2	553	47	AU034075	AU034075	
294	37	1.2	400	31	AA824581	oc83c02.s	AA824581	C 367	37	1.2	563	32	AA898161	NCM4H9T7	
295	37	1.2	400	35	AI131746	ue96a01.x	AI131746	C 368	37	1.2	571	63	AW067355	683022G06	
296	37	1.2	401	80	AW733536	sk74h03.y	AW733536	C 369	37	1.2	574	47	AU032687	AU032687	
297	37	1.2	404	42	AI624315	ts67c05.x	AI624315	C 370	37	1.2	574	48	AU083377	AU083377	
298	37	1.2	404	80	AW706146	AW706146	AW706146	C 371	37	1.2	575	39	AI401302	AI401302	
299	37	1.2	406	23	AA233530	zr30g06.s	AA233530	C 372	37	1.2	577	27	AA522882	n141b03.s	
300	37	1.2	409	41	AI579868	1028 NC-1	AI579868	C 373	37	1.2	580	119	A2066515	rpCI-23-4	
301	37	1.2	409	45	AI888650	wn22d07.x	AI888650	C 374	37	1.2	582	46	AI958190	fc91h10.y	
302	37	1.2	410	31	AA805093	CA89C12.s	AA805093	C 375	37	1.2	583	32	AA901876	NCC4C8T7	
303	37	1.2	410	35	AI090233	qb23b10.x	AI090233	C 376	37	1.2	588	43	AI730627	BNLGH1742	
304	37	1.2	411	44	AI762931	wh94e08.x	AI762931	C 377	37	1.2	597	30	AA701808	PMIFC_156	
305	37	1.2	417	79	AW633129	bl04b02.x	AW633129	C 378	37	1.2	603	47	AU032494	AU032494	
306	37	1.2	418	37	AI255423	u194d02.x	AI255423	C 379	37	1.2	604	36	AI207574	HA2938 Hu	
307	37	1.2	418	37	AI303093	ui74g01.x	AI303093	C 380	37	1.2	609	42	AI680453	AI680453	
308	37	1.2	422	27	AA489687	aa50c11.s	AA489687	C 381	37	1.2	610	48	AU038887	AU038887	
309	37	1.2	425	31	AA836303	od36d04.s	AA836303	C 382	37	1.2	611	36	AI172702	EST179 Ma	
310	37	1.2	429	28	AA569361	nh89f12.s	AA569361	C 383	37	1.2	611	36	AI174996	HA2736 Hu	
311	37	1.2	430	30	AA703338	zj11d06.s	AA703338	C 384	37	1.2	612	32	AA873373	Oh69b08.s	
312	37	1.2	430	81	C99609	C99609	C99609	C 385	37	1.2	614	64	AW108250	um20a02.x	
313	37	1.2	433	31	AA830841	oc54c01.s	AA830841	C 386	37	1.2	614	64	AI816377	au47c02.x	
314	37	1.2	436	62	AW019898	df01c10.y	AW019898	C 387	37	1.2	617	63	AW088253	xc99f01.x	
315	37	1.2	439	47	AL042490	DKFzp434G	AL042490	C 388	37	1.2	627	48	AU068669	AU068669	
316	37	1.2	448	36	AI160859	qb65f08.x	AI160859	C 389	37	1.2	632	27	AA504326	aa61e05.s	
317	37	1.2	448	39	AI431915	t126f10.x	AI431915	C 390	37	1.2	634	81	C96971	C96971	
318	37	1.2	448	69	AW188426	xj98f04.x	AW188426	C 391	37	1.2	646	64	AW164210	se23c10.y	
319	37	1.2	451	31	AA813963	OC06g07.s	AA813963	C 392	37	1.2	647	47	AU030767	AU030767	
320	37	1.2	451	47	AU029647	AU029647	AU029647	C 393	37	1.2	656	70	AW265928	L30-2458T	
321	37	1.2	452	31	AA815048	AA815048	AA815048	C 394	37	1.2	669	47	AL041134	DRFzP434F	
322	37	1.2	456	81	C91915	C91915	C91915	C 395	37	1.2	671	42	AI624433	ts29g03.x	
323	37	1.2	457	63	AW080747	AW080747	AW080747	C 396	37	1.2	676	80	AW725836	GA_Ea001	
324	37	1.2	458	48	AU057569	AU057569	AU057569	C 397	37	1.2	679	48	AU056612	AU056612	
325	37	1.2	458	79	AW632966	bl02c09.x	AW632966	C 398	37	1.2	684	39	AI395775	MA005127.	
326	37	1.2	465	39	AI440525	tc85g12.x	AI440525	C 399	37	1.2	712	48	AU075852	AU075852	
327	37	1.2	465	46	AI974961	EST269555	AI974961	C 400	37	1.2	718	40	AI478986	tn29d01.x	
328	37	1.2	467	48	AU057247	AU057247	AU057247	C 401	37	1.2	726	27	AA528799	nf52b07.s	
329	37	1.2	470	28	AA602351	np47g01.s	AA602351	C 402	37	1.2	726	81	C99640	C99640	
330	37	1.2	470	28	AU033739	AU033739	AU033739	C 403	37	1.2	729	62	AW019812	AW019812	
331	37	1.2	474	81	C98042	C98042	C98042	C 404	37	1.2	730	47	AL036742	DKFzP564I	
332	37	1.2	476	39	AI395813	MA005189.	AI395813	C 405	37	1.2	733	62	AV405504	AV405504	
333	37	1.2	476	44	AI829893	wj47a10.x	AI829893	C 406	37	1.2	991	122	CMS00JW8	AL077118 Drosophll	
334	37	1.2	483	23	AA277199	vb91g11.r	AA277199	C 407	36	1.2	49	27	AA526728	n191d10.s	
335	37	1.2	486	37	AI297124	LP11346.5	AI297124	C 408	36	1.2	49	28	AA608271	vn61f09.r	
336	37	1.2	486	47	AU029499	AU029499	AU029499	C 409	36	1.2	54	28	AA601314	nc15f06.s	
337	37	1.2	486	63	AW041017	EST283881	AW041017	C 410	36	1.2	54	42	AI633751	tt128b05.x	
338	37	1.2	487	47	AL045439	DKFzP434F	AL045439	C 411	36	1.2	60	72	AW409793	fn02d08.y	
339	37	1.2	489	48	AU057709	AU057709	AU057709	C 412	36	1.2	64	80	AW715605	g5g02nm.I	
340	37	1.2	489	79	AW642784	cm22e06.w	AW642784	C 413	36	1.2	69	29	AA639635	nq83e02.s	
341	37	1.2	494	32	AA898370	NCC2E1T7	AA898370	C 414	36	1.2	72	43	AI687854	ts89a03.x	
342	37	1.2	494	46	AI961384	wt17h04.x	AI961384	C 415	36	1.2	72	31	AA826924	OB52b03.s	
343	37	1.2	495	31	AA834558	od63d09.s	AA834558	C 416	36	1.2	79	89	T25558	EST00591 Un	
344	37	1.2	496	23	AA827427	pb07f09.r	AA827427	C 417	36	1.2	80	44	AI810694	tl19b09.x	
345	37	1.2	500	41	AI596140	uk25a02.x	AI596140	C 418	36	1.2	84	38	AI335218	ta93b09.x	
346	37	1.2	500	48	AU037533	AU037533	AU037533	C 419	36	1.2	84	89	AU053727	AU053727	
347	37	1.2	501	44	AI819127	wj79d02.x	AI819127	C 420	36	1.2	84	89	T25553	EST00586 Un	

c 421	36	1.2	85	28	AA606915	vm91c01.r	AA606915	vm91c01.r	c 494	36	1.2	141	30	AA720874	AA720874	nwl2h05.s
c 422	36	1.2	86	35	AI086378	oz44c01.x	AI086378	oz44c01.x	c 495	36	1.2	142	39	AI420269	AI420269	tf06c03.x
423	36	1.2	87	77	HSM001845	AI037514	Homo sapi	AI037514	496	36	1.2	142	62	AW022357	AW022357	df37d12.y
424	36	1.2	88	28	AA590663	vm24h10.r	AA590663	vm24h10.r	497	36	1.2	142	72	AW423982	AW423982	sh59a03.y
c 425	36	1.2	88	64	AW151979	xf71c01.x	AW151979	xf71c01.x	c 498	36	1.2	143	21	AA106528	AA106528	ml90a03.r
426	36	1.2	89	41	AI558863	co32e12.x	AI558863	co32e12.x	c 499	36	1.2	143	27	AA504514	AA504514	aa60h03.s
c 427	36	1.2	90	30	AW080028	W48671	zg45d06.s1	W48671	500	36	1.2	144	28	AA580663	AA580663	nd41b06.r
c 428	36	1.2	90	91	W48671	W48671	zg45d06.s1	W48671	c 501	36	1.2	144	31	AA814782	AA814782	ob42f09.s
c 429	36	1.2	91	41	AI543508	AI543508	tj23c08.x	AI543508	c 502	36	1.2	146	31	AA813232	AA813232	aj32g12.s
c 430	36	1.2	91	41	AI590755	AI590755	tw18d10.x	AI590755	c 503	36	1.2	146	39	AI437025	AI437025	fb37a02.y
431	36	1.2	93	46	AI960759	AI960759	sc90a09.y	AI960759	c 504	36	1.2	146	39	AI444985	AI444985	ti93d07.x
432	36	1.2	95	79	AW683162	AW683162	NFO08C12L	AW683162	c 505	36	1.2	146	41	AI559732	AI559732	co30a05.x
c 433	36	1.2	97	38	AI321720	AI321720	e2d07nm.f	AI321720	c 506	36	1.2	147	79	AW632958	AW632958	bl02c01.x
c 434	36	1.2	99	20	AA026667	AA026667	te93h01.s	AA026667	c 507	36	1.2	149	34	AI017427	AI017427	ou96f06.x
c 435	36	1.2	99	43	AI689157	AI689157	tx92f10.x	AI689157	c 508	36	1.2	149	38	AI345608	AI345608	bu83e06.x
436	36	1.2	100	28	AA590470	AA590470	vm24c02.r	AA590470	c 509	36	1.2	149	62	AW020763	AW020763	df14b04.y
437	36	1.2	100	28	AA606796	AA606796	vm90f07.r	AA606796	c 510	36	1.2	149	72	AW417850	AW417850	55157.MAR
438	36	1.2	100	72	AW438968	AW438968	77621.MAR	AW438968	c 511	36	1.2	150	42	AI652921	AI652921	wb40g10.x
439	36	1.2	101	28	AA555145	AA555145	nl07g04.s	AA555145	c 512	36	1.2	150	47	AI039118	AI039118	DKF2p566J
c 440	36	1.2	101	39	AI415241	AI415241	mdl4d03.x	AI415241	c 513	36	1.2	150	63	AW101452	AW101452	sd78d07.y
c 441	36	1.2	102	79	AW681730	AW681730	EST00448	AW681730	c 514	36	1.2	151	42	AI620145	AI620145	tu46b02.x
442	36	1.2	102	80	C21335	HUMGS000333	C21335	HUMGS000333	c 515	36	1.2	151	44	AI801558	AI801558	to91a03.x
443	36	1.2	103	47	AL048499	AL048499	DKS2p586M	AL048499	c 516	36	1.2	151	64	AW163834	AW163834	au86h07.y
444	36	1.2	104	48	AW053720	AW053720	AU053720	AU053720	c 517	36	1.2	152	28	AA582366	AA582366	nn53e01.s
445	36	1.2	105	22	AA216932	AA216932	mv75c12.r	AA216932	c 518	36	1.2	152	72	AW424113	AW424113	sh61a12.y
446	36	1.2	105	86	H89138	H89138	yw27a08.r1	H89138	c 519	36	1.2	153	41	AI590235	AI590235	to22c08.x
447	36	1.2	106	22	AA172938	AA172938	mt01g10.r	AA172938	c 520	36	1.2	153	81	C90066	C90066	C90066.Dict
448	36	1.2	106	72	AW416451	AW416451	51810.MAR	AW416451	c 521	36	1.2	154	21	AA145131	AA145131	mr08e01.r
c 449	36	1.2	108	48	AV021502	AV021502	AV021502	AV021502	c 522	36	1.2	154	70	AW315888	AW315888	13614.MAR
c 450	36	1.2	109	42	AI659651	AI659651	tu25b02.x	AI659651	c 523	36	1.2	154	79	AW633164	AW633164	b104e02.x
c 451	36	1.2	111	30	AA761557	AA761557	nz23c03.s	AA761557	c 524	36	1.2	155	23	AA242068	AA242068	mw23h12.r
c 452	36	1.2	111	37	AI289394	AI289394	qw21h08.x	AI289394	c 525	36	1.2	155	37	AI280010	AI280010	qu65h04.x
453	36	1.2	111	38	AI345471	AI345471	tb82f09.x	AI345471	c 526	36	1.2	155	46	AI938906	AI938906	sc62f01.y
c 454	36	1.2	112	42	AI612446	AI612446	486085C04	AI612446	c 527	36	1.2	155	73	AW479501	AW479501	25665.MAR
455	36	1.2	112	70	AW307653	AW307653	1387.MAR	AW307653	c 528	36	1.2	156	41	AI561356	AI561356	tq48e08.x
456	36	1.2	113	64	AW161202	AW161202	au79g06.y	AW161202	c 529	36	1.2	156	46	AI930742	AI930742	sc46d05.y
457	36	1.2	113	72	AW430509	AW430509	70135.MAR	AW430509	c 530	36	1.2	156	48	AU037919	AU037919	AU037919
458	36	1.2	113	77	HSM002662	AI038319	Homo sapi	AI038319	c 531	36	1.2	156	69	AW213895	AW213895	au44f05.x
c 459	36	1.2	114	39	AI433008	AI433008	th42e09.x	AI433008	c 532	36	1.2	156	73	AW479507	AW479507	25689.MAR
c 460	36	1.2	117	37	AI289791	AI289791	qw12c03.x	AI289791	c 533	36	1.2	157	45	AI868163	AI868163	to97b12.x
c 461	36	1.2	117	73	AW486832	AW486832	77218.MAR	AW486832	c 534	36	1.2	157	64	AW164657	AW164657	se75h04.y
c 462	36	1.2	118	33	AA939202	AA939202	oi99h05.s	AA939202	c 535	36	1.2	157	73	AW485902	AW485902	68596.MAR
c 463	36	1.2	118	39	AI444981	AI444981	ti93d01.x	AI444981	c 536	36	1.2	157	80	C23819	C23819	D1ct
c 464	36	1.2	119	63	AW089152	AW089152	xc79g07.x	AW089152	c 537	36	1.2	158	37	AI289386	AI289386	qw21g08.x
c 465	36	1.2	119	72	AW426534	AW426534	60821.MAR	AW426534	c 538	36	1.2	159	31	AA811586	AA811586	ob45e12.s
c 466	36	1.2	119	79	AW690406	AW690406	NFO32G12S	AW690406	c 539	36	1.2	159	39	AI394443	AI394443	tf79c07.x
c 467	36	1.2	120	70	AW264380	AW264380	nr02e06.x	AW264380	c 540	36	1.2	159	42	AI559504	AI559504	tq44g09.x
c 468	36	1.2	121	31	AA830333	AA830333	oc50b09.s	AA830333	c 541	36	1.2	159	64	AW151136	AW151136	xg33h03.x
c 469	36	1.2	121	63	AW065627	AW065627	614056D11	AW065627	c 542	36	1.2	159	64	AW151136	AW151136	xg33h03.x
c 470	36	1.2	121	72	AW436766	AW436766	77229.MAR	AW436766	c 543	36	1.2	159	71	AW336889	AW336889	21204.MAR
c 471	36	1.2	122	22	AA149757	AA149757	zo01f01.s	AA149757	c 544	36	1.2	159	81	C99492	C99492	C99492
c 472	36	1.2	123	79	AW681481	AW681481	EST000159	AW681481	c 545	36	1.2	160	22	AA151875	AA151875	zn99a08.s
c 473	36	1.2	125	38	AI361271	AI361271	qy42g06.x	AI361271	c 546	36	1.2	160	71	AW327210	AW327210	20914.MAR
c 474	36	1.2	125	79	AW642236	AW642236	cm17d03.w	AW642236	c 547	36	1.2	160	71	AW336921	AW336921	22554.MAR
c 475	36	1.2	127	63	AA028993	AA028993	wv98b09.x	AA028993	c 548	36	1.2	160	89	T04635	T04635	682.Lambda
c 476	36	1.2	128	23	AA227777	AA227777	mv99h10.r	AA227777	c 549	36	1.2	161	37	AI282595	AI282595	qv05a06.x
c 477	36	1.2	130	22	AA217439	AA217439	mu94a09.r	AA217439	c 550	36	1.2	161	64	AW118252	AW118252	xe12d12.x
c 478	36	1.2	130	42	AI611647	AI611647	tt37g05.x	AI611647	c 551	36	1.2	161	72	AW416297	AW416297	51510.MAR
c 479	36	1.2	130	48	AU037699	AU037699	AU037699	AU037699	c 552	36	1.2	162	79	AW416297	AW416297	ti95a09.x
c 480	36	1.2	131	33	AA933767	AA933767	om59h09.s	AA933767	c 553	36	1.2	162	47	AU032238	AU032238	se12d03.y
c 481	36	1.2	131	48	AU037779	AU037779	AU037779	AU037779	c 554	36	1.2	162	64	AW132939	AW132939	se12d03.y
c 482	36	1.2	131	77	HSM006247	AI041771	Homo sapi	AI041771	c 555	36	1.2	163	44	AI802244	AI802244	tj36e07.x
c 483	36	1.2	132	43	AI702797	AI702797	we07e02.x	AI702797	c 556	36	1.2	164	42	AI613028	AI613028	ty09g04.x
c 484	36	1.2	135	30	AA737267	AA737267	ao20e05.s	AA737267	c 557	36	1.2	164	43	AI758547	AI758547	xb21f09.x
c 485	36	1.2	135	41	AI560511	AI560511	tq46g06.x	AI560511	c 558	36	1.2	164	63	AW079012	AW079012	C92873
c 486	36	1.2	135	42	AI632096	AI632096	ts79g12.x	AI632096	c 559	36	1.2	164	81	C92873	C92873	D1ct
c 487	36	1.2	136	37	AI271763	AI271763	qj77d05.x	AI271763	c 560	36	1.2	165	30	AA766539	AA766539	oa33f03.s
c 488	36	1.2	136	72	AW429931	AW429931	68536.MAR	AW429931	c 561	36	1.2	165	40	AI471325	AI471325	tm09h02.x
489	36	1.2	136	72	AW432459	AW432459	sh74c12.y	AW432459	c 562	36	1.2	166	71	AW396791	AW396791	sg53a01.y
490	36	1.2	137	72	AA429406	AA429406	67718.MAR	AA429406	c 563	36	1.2	166	73	AW485427	AW485427	66369.MAR
c 491	36	1.2	140	32	AA878532	AA878532	oj19d09.s	AA878532	c 564	36	1.2	166	74	AW570333	AW570333	sj22f04.y
c 492	36	1.2	140	38	AI364783	AI364783	qul18e12.x	AI364783	c 565	36	1.2	167	20	AA024151	AA024151	mh97d04.r
493	36	1.2	140	48	AU056957	AU056957	AU056957	AU056957	c 566	36	1.2	167	30	AA749425	AA749425	ny12e01.s

567	36	1.2	167	47	AL079802	DKF2p4340	c 640	36	1.2	190	64	AW167460
568	36	1.2	168	22	AA189350	mt85c01.r	641	36	1.2	190	72	AW530315
569	36	1.2	168	23	AA276966	vc42h11.r	642	36	1.2	190	74	AW573456
570	36	1.2	168	42	AW428528	t253f09.x	643	36	1.2	191	37	AL044105
571	36	1.2	168	72	AW428528	66711.MAR	644	36	1.2	191	47	AL041772
572	36	1.2	169	36	AA1167912	ok29b07.x	645	36	1.2	191	74	AW569034
573	36	1.2	169	38	AA1345010	ta95b02.x	646	36	1.2	192	32	AA916898
574	36	1.2	169	41	AA1545595	fb67b10.x	647	36	1.2	192	37	AA175178
575	36	1.2	169	42	AA117971	DKF2p761F	648	36	1.2	192	37	AA1282031
576	36	1.2	170	21	AA116459	mp95d12.r	649	36	1.2	192	42	AA1622052
577	36	1.2	171	47	AL040077	DKF2p434A	650	36	1.2	192	45	AA1856699
578	36	1.2	172	29	AA638609	vo55c04.r	651	36	1.2	193	23	AA266283
579	36	1.2	172	42	AA1651093	wa97b11.x	652	36	1.2	193	27	AA490318
580	36	1.2	172	63	AA100667	sd8d11.Y	653	36	1.2	193	30	AA270850
581	36	1.2	172	71	AAW346438	28052.MAR	654	36	1.2	193	31	AA811663
582	36	1.2	173	35	AA1105939	cb03h10.t	655	36	1.2	193	35	AA1144106
583	36	1.2	173	41	AA1583023	tr99d10.x	656	36	1.2	193	37	AA1280528
584	36	1.2	173	47	AA120706	DKF2p762G	657	36	1.2	193	64	AA168700
585	36	1.2	173	64	AA1644430	se72d09.Y	658	36	1.2	193	117	AA922522
586	36	1.2	174	31	AA836253	od17c12.S	659	36	1.2	194	22	AA147918
587	36	1.2	174	41	AA1606084	ve97g04.x	660	36	1.2	194	29	AA693354
588	36	1.2	174	62	AAW020529	df11b04.Y	661	36	1.2	194	37	AA1285817
589	36	1.2	174	63	AAW082210	xb52a09.x	662	36	1.2	194	40	AA171278
590	36	1.2	174	72	AAW17645	54545.MAR	663	36	1.2	195	36	AA174394
591	36	1.2	174	72	AAW430301	69657.MAR	664	36	1.2	195	41	AA1540416
592	36	1.2	176	40	AA749184	va91e06.S	665	36	1.2	195	43	AA1689340
593	36	1.2	176	42	AA1661526	ny49g02.x	666	36	1.2	195	44	AA1824576
594	36	1.2	176	46	AA1952080	wx46g01.x	667	36	1.2	196	22	AA173255
595	36	1.2	176	63	AAW102540	sd60e12.Y	668	36	1.2	196	35	AA1105952
596	36	1.2	177	30	AA764903	nz76h11.S	669	36	1.2	196	81	FO0107
597	36	1.2	177	36	AA768725	oc86g07.S	670	36	1.2	196	89	TA03036
598	36	1.2	177	46	AA1934065	wn96b01.x	671	36	1.2	197	38	AA1376494
599	36	1.2	177	72	AAW396827	sg53g02.Y	672	36	1.2	197	45	AA189872
600	36	1.2	177	72	AAW414476	48167.MAR	673	36	1.2	198	22	AA189279
601	36	1.2	178	25	AA389924	vb28g06.r	674	36	1.2	198	47	AA134242
602	36	1.2	178	31	AA806857	oc29e05.S	675	36	1.2	198	63	AAW101681
603	36	1.2	178	47	AA273583	AJ273583	676	36	1.2	199	30	AA173657
604	36	1.2	179	20	AA009346	TgESTzz08	677	36	1.2	199	44	AA1796143
605	36	1.2	179	30	AAW67252	oc89f08.S	678	36	1.2	199	63	AAW087938
606	36	1.2	180	37	AA1241824	qu72d09.x	679	36	1.2	199	71	AAW654681
607	36	1.2	180	63	AAW089293	xd03e07.x	680	36	1.2	200	31	AA837290
608	36	1.2	181	39	AA1401088	th2zc02.x	681	36	1.2	200	32	AA911452
609	36	1.2	181	43	AA1701868	tu19a03.x	682	36	1.2	201	36	AA1232980
610	36	1.2	181	69	AAW227579	up11b08.x	683	36	1.2	201	37	AA1284131
611	36	1.2	181	73	AAW477996	18034.MAR	684	36	1.2	201	45	AA1890451
612	36	1.2	182	45	AA1866201	wn17h08.x	685	36	1.2	201	63	AAW103317
613	36	1.2	182	47	AA1119863	DKF2p761L	686	36	1.2	201	64	AAW168341
614	36	1.2	182	64	AAW156216	se21a09.Y	687	36	1.2	201	72	AAW423389
615	36	1.2	182	73	AAW79525	25756.MAR	688	36	1.2	202	22	AA198860
616	36	1.2	183	22	AA179931	zp56h09.S	689	36	1.2	202	23	AA274632
617	36	1.2	183	23	AA265647	mw14c11.r	690	36	1.2	203	23	AA275299
618	36	1.2	183	32	AA860080	ak45c09.S	691	36	1.2	203	23	AA275355
619	36	1.2	183	63	AAW029009	wv98d11.x	692	36	1.2	203	63	AAW100931
620	36	1.2	183	72	AAW47056	88151.MAR	693	36	1.2	204	23	AA289196
621	36	1.2	184	91	W32269	zc04h06.s1	694	36	1.2	204	40	AA193133
622	36	1.2	184	29	AA632400	np67f09.S	695	36	1.2	204	44	AA1768459
623	36	1.2	184	30	AA710574	vt43c11.r	696	36	1.2	204	63	AAW055261
624	36	1.2	184	31	AA805708	ns42c02.S	697	36	1.2	205	23	AA251340
625	36	1.2	184	63	AA102497	sd60a02.Y	698	36	1.2	205	81	C90715
626	36	1.2	185	37	AA1252901	qv31d08.x	699	36	1.2	206	21	AA125000
627	36	1.2	185	71	AAW36628	21730.MAR	700	36	1.2	206	48	AAU038875
628	36	1.2	186	23	AA258156	zs35a12.S	701	36	1.2	206	73	AAW483853
629	36	1.2	186	80	AAW178326	14g11nm.f	702	36	1.2	207	31	AAW795017
630	36	1.2	187	21	AA145886	mr33b12.r	703	36	1.2	207	37	AA1277260
631	36	1.2	187	28	AA611254	vo51a12.r	704	36	1.2	207	38	AA1284542
632	36	1.2	187	42	AA1620042	ty47f08.x	705	36	1.2	207	38	AA1352497
633	36	1.2	187	47	AA1044623	DKF2p4340	706	36	1.2	207	63	AAW101865
634	36	1.2	187	63	AAW099176	sd35f12.Y	707	36	1.2	207	80	AAW734306
635	36	1.2	187	64	AAW129264	xf22f05.x	708	36	1.2	208	27	AA512135
636	36	1.2	188	70	AAW07553	sf58g04.Y	709	36	1.2	208	33	AAU074644
637	36	1.2	188	73	AAW274350	TgESTzz25	710	36	1.2	208	34	AA1024643
638	36	1.2	189	47	AA1036214	DKF2p564H	711	36	1.2	208	35	AA1110737
639	36	1.2	190	33	AA959780	wv55b09.S	712	36	1.2	208	37	AA1266652

c 713	36	1.2	208	40	A1522772	A1522772	fb62c01.x	c 786	36	1.2	222	28	AA583860	AA583860	nn60c12.s
c 714	36	1.2	208	42	A1640130	A1640130	wa29e02.x	c 787	36	1.2	222	39	AI404030	AI404030	tj03a05.x
c 715	36	1.2	209	41	A1571909	A1571909	tr71b02.x	c 788	36	1.2	222	64	AW132524	AW132524	se05b11.y
c 716	36	1.2	209	43	A1683776	A1683776	tw53h07.x	c 789	36	1.2	222	72	AW425645	AW425645	57892.MAR
c 717	36	1.2	209	48	AU074565	AU074565	AU074565	c 790	36	1.2	223	37	AI250342	AI250342	qx04b01.x
c 718	36	1.2	209	63	AW102550	AW102550	sd60f12.y	c 791	36	1.2	224	33	AA936133	AA936133	om57c08.s
c 719	36	1.2	209	69	AW193280	AW193280	x172h08.x	c 792	36	1.2	224	41	AI155951	AI155951	tq50d07.x
c 720	36	1.2	209	71	AW345987	AW345987	26929.MAR	c 793	36	1.2	224	71	AW328117	AW328117	dt04d01.x
c 721	36	1.2	209	72	AW396227	AW396227	sh26a01.y	c 794	36	1.2	224	71	AW353417	AW353417	34219.MAR
c 722	36	1.2	209	72	AW426942	AW426942	61884.MAR	c 795	36	1.2	224	72	AW415603	AW415603	49994.MAR
c 723	36	1.2	209	72	AW429927	AW429927	58526.MAR	c 796	36	1.2	225	30	AA714186	AA714186	nw05g02.s
c 724	36	1.2	210	23	AA235975	AA235975	z805e01.s	c 797	36	1.2	225	37	AI289503	AI289503	qw28e04.x
c 725	36	1.2	210	28	AA587817	AA587817	rj06h02.s	c 798	36	1.2	225	47	AL040459	AL040459	DXF2p434D
c 726	36	1.2	210	71	AW396110	AW396110	sh25a04.y	c 799	36	1.2	225	64	AW168497	AW168497	DXF2p5861
c 727	36	1.2	211	119	AZ078723	AZ078723	RPCI-23-4	c 800	36	1.2	225	64	AW168497	AW168497	x186a07.x
c 728	36	1.2	212	26	AA411768	AA411768	zu02c04.s	c 801	36	1.2	225	70	AW307246	AW307246	sf54f03.y
c 729	36	1.2	212	27	AA489198	AA489198	aa57b01.s	c 802	36	1.2	225	70	AA29704	AA29704	68116.MAR
c 730	36	1.2	212	40	AI491817	AI491817	tn95a06.x	c 803	36	1.2	226	22	AA174678	AA174678	ms73h01.r
c 731	36	1.2	212	43	AI689557	AI689557	tx84c07.x	c 804	36	1.2	226	40	AI499446	AI499446	tn95e07.x
c 732	36	1.2	212	63	AW101775	AW101775	sd70b04.y	c 805	36	1.2	226	63	AW089171	AW089171	xc80h05.x
c 733	36	1.2	212	64	AW168563	AW168563	x189b03.x	c 806	36	1.2	226	63	AW090446	AW090446	xc83b09.x
c 734	36	1.2	213	20	AA038775	AA038775	mi95f04.r	c 807	36	1.2	226	64	AW151664	AW151664	xf67d10.x
c 735	36	1.2	213	32	AA878812	AA878812	cf83b03.s	c 808	36	1.2	226	71	AW353085	AW353085	35088.MAR
c 736	36	1.2	213	34	AI025427	AI025427	ow27g06.s	c 809	36	1.2	226	72	AA458709	AA458709	sh12f05.y
c 737	36	1.2	213	37	AI245325	AI245325	qk28e01.x	c 810	36	1.2	226	73	AA458709	AA458709	68259.MAR
c 738	36	1.2	213	45	AI1887241	AI1887241	vm37g08.x	c 811	36	1.2	227	32	AA937752	AA937752	o15g09.s
c 739	36	1.2	213	69	AW195444	AW195444	xn38d08.x	c 812	36	1.2	227	42	AI648663	AI648663	ty63g10.x
c 740	36	1.2	214	26	AA450649	AA450649	vf80c08.r	c 813	36	1.2	227	64	AW170775	AW170775	xj24g11.x
c 741	36	1.2	214	43	AI736553	AI736553	sb30c09.y	c 814	36	1.2	227	64	AW170775	AW170775	xj24g11.x
c 742	36	1.2	214	48	AU053098	AU053098	se36c12.y	c 815	36	1.2	228	31	AA833731	AA833731	aj48c01.s
c 743	36	1.2	214	64	AW153129	AW153129	vm50h12.y	c 816	36	1.2	228	70	AW263347	AW263347	qx87h07.x
c 744	36	1.2	215	45	AI874303	AI874303	vm50h12.y	c 817	36	1.2	228	103	AQ487403	AQ487403	RPCI-11-2
c 745	36	1.2	216	20	AA047715	AA047715	zf15a01.r	c 818	36	1.2	229	29	AA633592	AA633592	ac15f07.s
c 746	36	1.2	216	42	AI628651	AI628651	ty77e12.x	c 819	36	1.2	229	30	AA711151	AA711151	vt56e08.r
c 747	36	1.2	216	70	AW314680	AW314680	11001.MAR	c 820	36	1.2	229	41	AI583152	AI583152	tr98e04.x
c 748	36	1.2	216	71	AW344655	AW344655	26750.MAR	c 821	36	1.2	229	42	AI810736	AI810736	tu04a07.x
c 749	36	1.2	216	72	AW429999	AW429999	68710.MAR	c 822	36	1.2	229	72	AA459376	AA459376	sh41c04.y
c 750	36	1.2	217	28	AA596229	AA596229	vo30h01.r	c 823	36	1.2	230	20	AA020512	AA020512	mb61f01.r
c 751	36	1.2	217	32	AA873481	AA873481	oh77g03.s	c 824	36	1.2	230	43	AI736133	AI736133	8b23d05.y
c 752	36	1.2	217	37	AI272994	AI272994	qv63e01.x	c 825	36	1.2	230	63	AW101750	AW101750	sg69g09.y
c 753	36	1.2	217	45	AI660051	AI660051	tx72h04.x	c 826	36	1.2	230	80	C23731	C23731	D1c1
c 754	36	1.2	217	48	AU078401	AU078401	AU078401	c 827	36	1.2	231	231	C23731	C23731	D1c1
c 755	36	1.2	217	79	AW631928	AW631928	91315.MAR	c 828	36	1.2	231	32	AI252030	AI252030	qv39c05.x
c 756	36	1.2	218	30	AA713785	AA713785	nv70b07.s	c 829	36	1.2	231	42	AI636811	AI636811	ts60a06.x
c 757	36	1.2	218	38	AA766268	AA766268	oa29b05.s	c 830	36	1.2	231	47	AT000896	AT000896	AT000896
c 758	36	1.2	218	38	AI351164	AI351164	ql11h12.x	c 831	36	1.2	231	90	WI4959	WI4959	md33e04.r1
c 759	36	1.2	218	47	AL045986	AL045986	DKF2p4340	c 832	36	1.2	232	30	AA743445	AA743445	ny19a10.s
c 760	36	1.2	218	62	AU007733	AU007733	wt68e01.x	c 833	36	1.2	232	48	AW073164	AW073164	AU073164
c 761	36	1.2	218	62	AU007938	AU007938	vw47a07.x	c 834	36	1.2	232	70	AW306696	AW306696	sf47b10.y
c 762	36	1.2	218	63	AW099572	AW099572	sd43f11.y	c 835	36	1.2	233	43	AI887752	AI887752	vm18c09.x
c 763	36	1.2	218	72	AW459746	AW459746	sh91e09.y	c 836	36	1.2	233	63	AW057680	AW057680	wx01g07.x
c 764	36	1.2	219	20	AA027470	AA027470	mi05d01.r	c 837	36	1.2	233	88	R65859	R65859	y123a08.r1
c 765	36	1.2	219	30	AA709719	AA709719	vt38a10.r	c 838	36	1.2	234	23	AA321147	AA321147	mw12g01.r
c 766	36	1.2	219	39	AI452993	AI452993	tj46c09.x	c 839	36	1.2	234	39	AI387007	AI387007	GH17416.3
c 767	36	1.2	219	45	AI865070	AI865070	wk07b02.x	c 840	36	1.2	234	47	AL043073	AL043073	DXF2p434B
c 768	36	1.2	219	64	AW156260	AW156260	se21f03.y	c 841	36	1.2	234	63	AW100583	AW100583	sq57c02.y
c 769	36	1.2	219	64	AW156522	AW156522	se28c01.y	c 842	36	1.2	234	103	AQ483178	AQ483178	RPCI-11-2
c 770	36	1.2	219	70	AW311182	AW311182	sg35e12.y	c 843	36	1.2	235	85	F34689	F34689	HSPD29912.H
c 771	36	1.2	219	85	F37319	F37319	HSPD35559.H	c 844	36	1.2	236	25	AA239005	AA239005	mx95g02.r
c 772	36	1.2	219	86	H84469	H84469	yv66f12.r1	c 845	36	1.2	236	40	AI521077	AI521077	tc70f05.x
c 773	36	1.2	220	70	AW312421	AW312421	4218.MAR	c 846	36	1.2	236	42	AI678494	AI678494	tu83b04.x
c 774	36	1.2	221	23	AA270235	AA270235	va65a07.r	c 847	36	1.2	236	63	AW087987	AW087987	xb45h06.x
c 775	36	1.2	221	32	AA918063	AA918063	ol81a12.s	c 848	36	1.2	237	22	AA192029	AA192029	rs04g02.r
c 776	36	1.2	221	35	AI144071	AI144071	ql63e07.x	c 849	36	1.2	237	23	AA269780	AA269780	va55b06.r
c 777	36	1.2	221	36	AI203594	AI203594	qf57g03.x	c 850	36	1.2	237	42	AI653303	AI653303	wb24a04.x
c 778	36	1.2	221	39	AI414351	AI414351	mb64e11.x	c 851	36	1.2	238	21	AA123988	AA123988	mq21d11.r
c 779	36	1.2	221	41	AI569953	AI569953	tr90c02.x	c 852	36	1.2	238	39	AI451648	AI451648	mu55e09.x
c 780	36	1.2	221	42	AI643858	AI643858	vu10d11.x	c 853	36	1.2	238	41	AI564988	AI564988	tc45b07.x
c 781	36	1.2	221	79	AW659288	AW659288	96321.MAR	c 854	36	1.2	238	41	AI641467	AI641467	fc16a05.x
c 782	36	1.2	222	86	H83576	H83576	yv81a04.r1	c 855	36	1.2	238	64	AQ078631	AQ078631	CTT-HSP-2
c 783	36	1.2	222	86	AA017921	AA017921	mh46b08.r	c 856	36	1.2	238	93	AQ078631	AQ078631	CTT-HSP-2
c 784	36	1.2	222	22	AA189894	AA189894	mu55c12.r	c 857	36	1.2	239	37	AI272328	AI272328	ap21g10.x
c 785	36	1.2	222	23	AA258131	AA258131	zs35f04.s	c 858	36	1.2	239	39	AI452772	AI452772	tj45d02.x

859	36	1.2	239	71	AW325910	AW325910 16971 MAR	c 932	36	1.2	252	42	AI675339	AI675339 wb98b09.x
860	36	1.2	240	20	AA021163	AA021163 ze6se10.s	c 933	36	1.2	252	45	AI872711	AI872711 wm69c04.x
861	36	1.2	240	27	AA519067	AA519067 v126h07.r	934	36	1.2	252	48	AU070434	AU070434 AU070434
862	36	1.2	240	47	AL043632	AL043632 DKFP434I	935	36	1.2	252	70	AW314957	AW314957 11566 MAR
863	36	1.2	240	48	AU073934	AU073934 AU073934	936	36	1.2	253	21	AA138875	AA138875 mt-03q07.r
864	36	1.2	241	22	AA204591	AA204591 mu5sail1.r	c 937	36	1.2	253	42	AI642148	AI642148 uc46d12.x
865	36	1.2	241	62	AW020414	AW020414 df09d09.y	938	36	1.2	253	70	AW313997	AW313997 9738 MARC
866	36	1.2	241	63	AW101194	AW101194 sg75c08.y	939	36	1.2	253	72	AW423296	AW423296 sh65q05.y
867	36	1.2	241	70	AW312799	AW312799 5109 MARC	940	36	1.2	254	22	AA171990	AA171990 z998h05.s
868	36	1.2	241	120	B51917	B51917 CJT-HSP-384	c 941	36	1.2	254	38	AI309193	AI309193 q070b07.x
869	36	1.2	242	27	AA472916	AA472916 vb31c02.r	942	36	1.2	254	40	AI529455	AI529455 v443401.y
870	36	1.2	242	28	AA583602	AA583602 nn61e05.s	943	36	1.2	254	48	AU068379	AU068379 AU068379
871	36	1.2	242	34	AI011695	AI011695 EST206146	944	36	1.2	254	48	AU077955	AU077955 AU077955
872	36	1.2	242	38	AI345463	AI345463 lb82f01.x	945	36	1.2	254	70	AW314835	AW314835 11211 MAR
873	36	1.2	242	38	AI382313	AI382313 ta71c08.x	946	36	1.2	254	72	AW426311	AW426311 60099 MAR
874	36	1.2	242	63	AW102544	AW102544 sg60f05.y	947	36	1.2	254	73	AW482472	AW482472 44642 MAR
875	36	1.2	242	63	AW105459	AW105459 xd53a05.x	948	36	1.2	255	34	AI053520	AI053520 q168d12.x
876	36	1.2	242	79	AW656544	AW656544 108855 MA	c 949	36	1.2	255	37	AI249880	AI249880 qx03609.x
877	36	1.2	243	26	AA4011240	AA4011240 zv63q02.r	950	36	1.2	255	72	AW432895	AW432895 sh99d12.y
878	36	1.2	243	44	AI766785	AI766785 w188c10.x	951	36	1.2	255	72	AW458259	AW458259 sh80d10.y
879	36	1.2	244	31	AA805575	AA805575 ob44g03.s	952	36	1.2	255	79	AW656569	AW656569 10891 MA
880	36	1.2	244	37	AI2511769	AI2511769 qu76e02.x	953	36	1.2	256	22	AA198975	AA198975 mv39a08.r
881	36	1.2	244	41	AI545287	AI545287 fb80c03.x	954	36	1.2	256	23	AA275331	AA275331 vc06g11.r
882	36	1.2	244	63	AW082599	AW082599 xc20b09.x	c 955	36	1.2	256	31	AA775242	AA775242 ad18a07.s
883	36	1.2	244	79	AW620896	AW620896 sj50a12.y	c 956	36	1.2	256	46	AI951062	AI951062 wx63a05.x
884	36	1.2	244	80	AW733574	AW733574 sk75d03.y	957	36	1.2	257	79	AW671911	AW671911 LG1_352_B
885	36	1.2	244	85	H42271	H42271 yo63b07.r1	958	36	1.2	258	22	AA189702	AA189702 mc90f03.r
886	36	1.2	245	30	AA761917	AA761917 rz24h03.s	959	36	1.2	258	23	AA260471	AA260471 v494e02.r
887	36	1.2	245	37	AI289992	AI289992 qw25e11.x	c 960	36	1.2	258	27	AA516371	AA516371 nf59f11.s
888	36	1.2	245	70	AW306382	AW306382 se49h09.y	961	36	1.2	258	79	AW668950	AW668950 111650 MA
889	36	1.2	245	70	AW311640	AW311640 sg43d06.y	c 962	36	1.2	258	87	N33147	N33147 yy06f05.sl
890	36	1.2	245	73	AW485189	AW485189 64429 MAR	c 963	36	1.2	259	45	AI886119	AI886119 wn16b06.x
891	36	1.2	246	23	AA238442	AA238442 mx92h10.r	964	36	1.2	259	91	W35226	W35226 zc70f01.r1
892	36	1.2	246	23	AA268648	AA268648 va43c04.r	965	36	1.2	259	34	AI006957	AI006957 us77c12.r
893	36	1.2	246	30	AA710129	AA710129 vc45a02.r	966	36	1.2	260	72	AW415379	AW415379 49631 MAR
894	36	1.2	246	42	AI653918	AI653918 fy03c05.x	967	36	1.2	260	80	AW733896	AW733896 sh78g09.y
895	36	1.2	246	43	AI696588	AI696588 lx68b08.x	c 968	36	1.2	261	28	AA618452	AA618452 nn77f09.s
896	36	1.2	246	47	AU031785	AU031785 AU031785	c 969	36	1.2	261	29	AA648323	AA648323 ns20h08.s
897	36	1.2	246	70	AW258200	AW258200 ug31f10.y	c 970	36	1.2	261	37	AI252644	AI252644 qu66h02.x
898	36	1.2	246	71	AW336324	AW336324 23232 MAR	c 971	36	1.2	261	41	AI564906	AI564906 tm86f07.x
899	36	1.2	246	71	AW337090	AW337090 22233 MAR	972	36	1.2	261	46	AI946035	AI946035 ds19h08.y
900	36	1.2	247	33	AA936138	AA936138 cm57d02.s	973	36	1.2	261	69	AW209723	AW209723 u146d04.y
901	36	1.2	247	37	AI289937	AI289937 qv08a03.x	974	36	1.2	262	38	AA741149	AA741149 nz01g06.s
902	36	1.2	247	40	AI522809	AI522809 fb22f10.x	975	36	1.2	262	38	AI312975	AI312975 ts80c04.x
903	36	1.2	247	43	AI725206	AI725206 1225 Pt1F	c 976	36	1.2	262	64	AW118554	AW118554 xe79f01.x
904	36	1.2	247	48	AW058329	AW058329 AU058329	c 977	36	1.2	262	81	C99495	C99495 C99495 Rice
905	36	1.2	247	63	AW065234	AW065234 614038B07	978	36	1.2	262	90	W08305	W08305 mb40a10.r1
906	36	1.2	248	30	AA738097	AA738097 nx11h10.s	c 978	36	1.2	263	23	AA221252	AA221252 mw11f03.r
907	36	1.2	248	30	AA767127	AA767127 oa61e09.s	979	36	1.2	263	34	AI061101	AI061101 au26b08.x
908	36	1.2	248	39	AI445994	AI445994 tj32c09.x	c 980	36	1.2	263	41	AI538127	AI538127 ts83f03.x
909	36	1.2	248	39	AU033156	AU033156 AU033156	c 981	36	1.2	262	64	AW118554	AW118554 xe79f01.x
910	36	1.2	248	72	AW414803	AW414803 48629 MAR	982	36	1.2	262	81	C99495	C99495 C99495 Rice
911	36	1.2	248	72	AA430974	AA430974 71107 MAR	983	36	1.2	262	90	W08305	W08305 mb40a10.r1
912	36	1.2	249	27	AA507581	AA507581 nh76g03.s	984	36	1.2	263	23	AA221252	AA221252 mw11f03.r
913	36	1.2	249	34	AI068033	AI068033 EST209722	c 985	36	1.2	263	34	AI061101	AI061101 au26b08.x
914	36	1.2	249	39	AI427562	AI427562 mm17e06.x	c 986	36	1.2	263	37	AI282827	AI282827 qt71c04.x
915	36	1.2	249	44	AI783562	AI783562 tz37c08.x	987	36	1.2	263	47	AU030162	AU030162 AU030162
916	36	1.2	249	70	AW311129	AW311129 sg33c11.y	988	36	1.2	264	47	AU031837	AU031837 AU031837
917	36	1.2	249	71	AW359405	AW359405 45979 MAR	c 989	36	1.2	264	64	AW151336	AW151336 xe73b09.x
918	36	1.2	249	72	AA415599	AA415599 49989 MAR	c 990	36	1.2	264	69	AW190401	AW190401 u162d07.x
919	36	1.2	249	72	AA427866	AA427866 64424 MAR	991	36	1.2	264	87	N81195	N81195 yw36h06.r1
920	36	1.2	250	22	AA147342	AA147342 zc039b03.s	c 992	36	1.2	265	22	AA159368	AA159368 z078h01.s
921	36	1.2	250	32	AA866865	AA866865 ud06c05.r	c 993	36	1.2	265	32	AA872519	AA872519 oa16b02.s
922	36	1.2	251	29	AA687139	AA687139 nv63e05.s	c 994	36	1.2	265	62	AW006533	AW006533 wt05h08.x
923	36	1.2	251	30	AA734985	AA734985 vs17g07.r	c 995	36	1.2	265	63	AW103608	AW103608 xs80h03.x
924	36	1.2	251	31	AA805063	AA805063 ob86a12.s	c 996	36	1.2	266	40	AI500634	AI500634 ts98h05.x
925	36	1.2	251	31	AA811218	AA811218 os77h07.s	997	36	1.2	266	64	AW132169	AW132169 sd67a07.y
926	36	1.2	251	32	AA877935	AA877935 of77h07.s	998	36	1.2	266	46	AI964444	AI964444 496010E08
927	36	1.2	251	38	AI344947	AI344947 tb01c03.x	999	36	1.2	268	47	AU031568	AU031568 AU031568
928	36	1.2	251	41	AI553332	AI553332 vw51f11.x	1000	36	1.2	268	47	AU032197	AU032197 AU032197
929	36	1.2	251	41	AI597465	AI597465 v164g11.x							
930	36	1.2	251	63	AW084330	AW084330 xc55c06.x							
931	36	1.2	252	23	AA230854	AA230854 mw14b02.r							

ALIGNMENTS


```

QY 818 ggaaggactggaagatcggttcacactgtgcaaccagacaacaaatcagactg 869
|||||
Db 257 GGAAGGACTGGAAGATCGGCTTCCAACTGTGCAACCAAGACCAACAAATCAGACTG 308
|||||

RESULT 5
AA980401/c
LOCUS
DEFINITION
ua52g11.r1 Stratagene mouse lung 937302 Mus musculus cDNA clone
IMAGE:1350404 5' similar to SW:SCAA_RAT P37089 AMILORIDE-SENSITIVE
SODIUM CHANNEL ALPHA-SUBUNIT ;, mRNA sequence.
ACCESSION
AA980401
VERSION
AA980401.1 GI:3158937
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 442)
REFERENCE
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMMI Mouse EST Project
Unpublished (1996)
On Apr 7, 1998 this sequence version replaced gi:3035617.
Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:699196
Possible reversed clone: similarity on wrong strand
Seq primer: -29ml3 rev1 ET from Amersham
High quality sequence stop: 439.
FEATURES
Location/Qualifiers
source
1..442
/organism="Mus musculus"
/strain="C57BL/6 x CBA"
/db_xref="taxon:10090"
/clone_lib="IMAGE:1350404"
/clone_id="Stratagene mouse lung 937302"
/sex="female"
/tissue_type="lung"
/dev_stage="6-8 month old"
/lab_host="SOLR (kanamycin resistant)"
/notes="Organ: lung; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. 6-8 month old female lung and 1.5 year old male
lung were source of mRNA. Average insert size: 1.5 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'
1 others"
BASE COUNT 112 a 114 c 150 g 65 t
ORIGIN

Query Match 1.4%; Score 45; DB 33; Length 442;
Best Local Similarity 100.0%; Pred. No. 8.8e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2061 gaccttccccctgacctgacagccctccacctgctatgctac 2105
|||||
Db 213 GACCCCTCCCTGGCCCTGACAGCCCTCCACCTGCCTATGCTAC 169
|||||

RESULT 6
AL040100/c
LOCUS
DEFINITION
AL040100.1
VERSION
AL040100.1 GI:5409067
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 157)
REFERENCE
Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Koehler, et al.)
Unpublished (1999)
On May 18, 1998 this sequence version replaced gi:3138571.
Contact: Koehler K
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
r1 sequence also available.
This clone (DKFZp434B0913) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
source
1..157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434B0913"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 38 a 14 c 25 g 79 t
ORIGIN

Query Match 1.3%; Score 41; DB 47; Length 157;
Best Local Similarity 100.0%; Pred. No. 7.8e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 54 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 14
|||||

RESULT 7
AW278153
LOCUS
DEFINITION
sf40408.y1 Gm-cl009 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl009-2440 5', mRNA sequence.
ACCESSION
AW278153
VERSION
AW278153.1 GI:6666694
KEYWORDS
EST.
SOURCE
soybean.
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales;
Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 157)
REFERENCE
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Corvelli, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
JOURNAL

```

COMMENT

On Oct 8, 1998 this sequence version replaced gi:3727461.
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134. For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or
 info@genomesystems.com web site: www.genomesystems.com
 Seq primer: -40RFP from Gibco
 High quality sequence stop: 132.

FEATURES

source
 1. .157
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl009-2440"
 /clone_lib="Gm-cl009"
 /lab_host="XL10-Gold"
 /note="vector: pBluescript II XR; Site_1: EcoRI; Site_2:
 XhoI; The mRNA was isolated from entire roots of
 2-month-old 'Williams' plants that were greenhouse grown
 in 5-gallon pots. To suppress nodulation, Black Gold
 All-Purpose potting soil was supplemented with: 0.36g/L
 available phosphoric acid (P205), 20mg/L urea N, 0.16g/L
 S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.53mg/L Mn,
 0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following
 nutrients in a slow-release form (Osmocote): 0.165g/L
 ammoniacal N, 0.185g/L nitrate N, 0.35g/L available
 phosphoric acid, and 0.35g/L soluble potash. No nodules
 were visible on the roots at harvest. Stratagene's cDNA
 Synthesis Kit (catalog #200401) was used to synthesize the
 cDNA. First-strand synthesis was performed with 5-methyl
 dCTP, hence the ligated cDNA is hemimethylated.
 Stratagene's first-strand synthesis primer was used
 [GACAGAGACAGAGAGAGAACTAGTCGAG(T)-18]. After
 second-strand synthesis is, the cDNA ends were 'polished'
 with clone Pfu DNA polymerase, ligated to EcoRI adapters,
 and phosphorylated. The XhoI site within the first-strand
 synthesis primer was restricted by digestion with XhoI;
 all XhoI sites in the cDNA would be protected by their
 hemimethylated status. The cDNA constructs were
 size-fractionated with a 400bp cutoff, using a SizeSep 400
 spin column from Pharmacia. The column eluent was then
 ligated into Stratagene's pBluescript II XR Predigested
 vector (pBluescript II SK(+)) that had been digested with
 EcoRI and XhoI, and phosphorylated. Both the white and
 blue colonies appear to contain recombinant plasmids with
 cDNA inserts. This library was constructed by Dr. Paul
 Keim and Dr. Virginia Coryell."

BASE COUNT

ORIGIN 99 a 21 c 12 g 25 t

Query Match 1.3%; Score 41; DB 70; Length 157;
 Best Local Similarity 100.0%; Pred. No. 7.8e-09;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

Db 69 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 109

RESULT 8

AU073712 240 bp mRNA EST 24-JUN-1999
 LOCUS AU073712 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
 DEFINITION AU073712 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
 accession AU073712
 VERSION AU073712.1 GI:5180133

KEYWORDS

SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum
 REFERENCE Eukaryote; Dictyostellida; Dictyostelium.
 AUTHORS Urushihara, H.
 TITLE Developmental cDNA in Dictyostelium discoideum (1999)
 JOURNAL Unpublished (1999)
 COMMENT On Jul 8, 1999 this sequence version replaced gi:5422501.
 Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
 Email: d402hu@sakura.cc.tsukuba.ac.jp
 PROJECT = Dictyostelium discoideum cDNA project in Japan.

FEATURES

source
 1. .240
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="SSI133"
 /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
 /dev_stage="slug" 30 g 85 t
 BASE COUNT 108 a 17 c 30 g 85 t
 ORIGIN

Query Match 1.3%; Score 41; DB 48; Length 240;
 Best Local Similarity 100.0%; Pred. No. 7.7e-09;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

Db 47 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 87

RESULT 9

AU065863 262 bp mRNA EST 30-NOV-1999
 LOCUS AU065863 Rice green shoot Oryza sativa cDNA clone SI4374_1A, mRNA
 DEFINITION sequence.
 ACCESSION AU065863
 VERSION AU065863.2 GI:6481240
 KEYWORDS EST.
 SOURCE Oryza sativa.

ORGANISM

Oryza sativa
 Eukaryote; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

REFERENCE

1 (bases 1 to 262)

AUTHORS

Yamamoto, K. and Sasaki, T.

TITLE

Rice cDNA from green shoot

JOURNAL

Unpublished (1996)

COMMENT

On Jun 2, 1999 this sequence version replaced gi:4968959.
 Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program
 2-1-2 Kannondai, Tsukuba
 Ibaraki,
 Japan 305
 Tel: 0298-38-7441
 Fax: 0298-38-7468
 Email: tsasaki@abr.affrc.go.jp
 PROJECT = 'RGP'
 Sequence updated (24-Nov-1999).

FEATURES

source
 1. .262
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /clone="SI4374_1A"
 /note="Green shoot (8 days old)"
 BASE COUNT 94 a 39 c 52 g 77 t

ORIGIN

Query Match 1.3%; Score 41; DB 48; Length 262;
 Best Local Similarity 100.0%; Pred. No. 7.7e-09; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 222 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 262

RESULT 10

AF034791 1367 bp mRNA EST 22-APR-1998
 LOCUS AF034791 Felis catus partial mRNA Felis catus cDNA similar to
 DEFINITION glyceraldehyde 3-phosphate dehydrogenase, mRNA sequence.
 ACCESSION AF034791
 VERSION AF034791.1 GI:2654165
 KEYWORDS EST.
 SOURCE cat.

ORGANISM

Felis catus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

REFERENCE

1 (bases 1 to 1367)
 Helps.C.R. and Harbour.D.A.

AUTHORS

Felis catus glyceraldehyde 3-phosphate dehydrogenase mRNA

JOURNAL

Unpublished (1997)

COMMENT

On Apr 7, 1998 this sequence version replaced gi:3036550.
 Contact: Helps, Chris R.
 Clinical and Veterinary Science
 University of Bristol
 Bristol BS18 7DU, UK
 Email: c.r.helps@bris.ac.uk.

FEATURES

source

1..1367
 /organism="Felis catus"
 /db_xref="taxon:9685"
 /clone_lib="Felis catus partial mRNA"
 /cell_line="FL4"

BASE COUNT 325 a 317 c 376 g 276 t 73 others

ORIGIN

Query Match 1.3%; Score 41; DB 33; Length 1367;
 Best Local Similarity 100.0%; Pred. No. 7.4e-09; Mismatches 0; Indels 0; Gaps 0;

Qy 3077 ttccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 1302 TTCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1342

RESULT 11

AW313563 210 bp mRNA EST 24-JAN-2000
 LOCUS AW313563 8840 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION AW313563
 ACCESSION AW313563
 VERSION AW313563.1 GI:6742783
 KEYWORDS EST.
 SOURCE Bos taurus.

ORGANISM

Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 210)

AUTHORS

Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
 Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and
 Keefe,J.W.

TITLE

Design and use of four pooled tissue normalized cDNA libraries for

JOURNAL

EST discovery in cattle

COMMENT

Unpublished (2000)
 On Jul 7, 1999 this sequence version replaced gi:5410190.

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACACGCTATGACCAT

BACKWARD: GTTTCCAGTCACGACG

Plate: 5 row: E column: 22

Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

1..210

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 1BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."

BASE COUNT 89 a 42 c 35 g 44 t

ORIGIN

Query Match 1.3%; Score 40; DB 70; Length 210;
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 tcctcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 157 TCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 196

RESULT 12

AW657118 224 bp mRNA EST 05-APR-2000
 LOCUS AW657118 109764 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION AW657118
 ACCESSION AW657118.1 GI:7422944
 VERSION
 KEYWORDS EST.
 SOURCE Bos taurus.

ORGANISM

Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 224)

AUTHORS

Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
 Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and
 Keefe,J.W.

TITLE

Design and use of four pooled tissue normalized cDNA libraries for

JOURNAL

EST discovery in cattle

COMMENT

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACACGCTATGACCAT

BACKWARD: GTTTCCAGTCACGACG

Plate: 99 row: H column: 21

Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

1..224

/organism="Bos taurus"

/db_xref="taxon:9913"
 /clone_lib="MARC IBOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."
 84 a 49 c 51 g 40 t

BASE COUNT
 ORIGIN

Query Match 1.3%; Score 40; DB 79; Length 224;
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 97 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 136

RESULT 13

LOCUS C98601 275 bp mRNA EST 19-OCT-1998
 DEFINITION C98601 Rice panicle at flowering stage Oryza sativa cDNA clone
 E0453_62, mRNA sequence.

ACCESSION C98601.1 GI:3761353
 VERSION C98601
 KEYWORDS EST.
 SOURCE Oryza sativa.

ORGANISM

REFERENCE 1 (bases 1 to 275)
 AUTHORS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 TITLE Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
 JOURNAL Sasaki.T. and Yamamoto.K.
 COMMENT Rice cDNA from panicle at flowering stage
 Unpublished (1996)
 On Jun 22, 1998 this sequence version replaced gi:3247562.

CONTACT: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program
 2-1-2 Kannondai, Tsukuba
 Ibaraki,
 Japan 305
 Tel: 0298-38-7441
 Fax: 0298-38-7468
 Email: tsasaki@abr.affrc.go.jp
 PROJECT -"RGP".

FEATURES

Location/Qualifiers
 1..275
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /clone_lib="E0453_62"
 /dev_stage="flowering stage"
 /note="Organ: panicle; Rice cDNA from panicle at flowering
 stage"
 113 a 44 c 42 g 70 t 6 others

BASE COUNT
 ORIGIN

Query Match 1.3%; Score 40; DB 81; Length 275;
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 225 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 264

RESULT 14

LOCUS AI253246 449 bp mRNA EST 05-NOV-1998

DEFINITION

ACCESSION AI253246
 VERSION AI253246.1 GI:3849775
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 449)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 unknown library type
 Seq primer: -40UP from Gibco
 High quality sequence stop: 436.

FEATURES
 Location/Qualifiers
 1..449

Source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2029294"
 /clone_lib="NCI_CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneds 1322376-1323911, 1456007-1456775, and
 1500552-1502855). Subtraction by Bento Soares and M.
 Fatima Bonaudo."

BASE COUNT 145 a 57 c 63 g 184 t
 ORIGIN

Query Match 1.3%; Score 40; DB 37; Length 449;
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 41 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2

RESULT 15

AI320733/c
 LOCUS AI320733 79 bp mRNA EST 18-DEC-1998
 DEFINITION c9f03nm.f1 Neurospora crassa morning cDNA library Neurospora crassa
 cDNA clone c9f03nm 5', mRNA sequence.

ACCESSION AI320733
 VERSION AI320733.1 GI:4036715
 KEYWORDS EST.
 SOURCE Neurospora crassa.
 ORGANISM Neurospora crassa
 Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae;
 Neurospora.

REFERENCE 1 (bases 1 to 79)
 AUTHORS Zhu.H., Lai.H., Kupfer.D., Dunlap,J.C. and Roe,B.A.
 TITLE Two Neurospora crassa EST Databases
 JOURNAL Unpublished (1998)
 COMMENT On May 18, 1998 this sequence version replaced gi:3138462.
 Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for genome technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762

qz39e12.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2029294 3',
 mRNA sequence.
 AI253246
 AI253246.1 GI:3849775
 EST.
 human.
 Homo sapiens
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 449)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 unknown library type
 Seq primer: -40UP from Gibco
 High quality sequence stop: 436.

FEATURES
 Location/Qualifiers
 1..449

Source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2029294"
 /clone_lib="NCI_CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneds 1322376-1323911, 1456007-1456775, and
 1500552-1502855). Subtraction by Bento Soares and M.
 Fatima Bonaudo."

BASE COUNT 145 a 57 c 63 g 184 t
 ORIGIN

Query Match 1.3%; Score 40; DB 37; Length 449;
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 tccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||
 Db 41 TCCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2

RESULT 15

AI320733/c
 LOCUS AI320733 79 bp mRNA EST 18-DEC-1998
 DEFINITION c9f03nm.f1 Neurospora crassa morning cDNA library Neurospora crassa
 cDNA clone c9f03nm 5', mRNA sequence.

ACCESSION AI320733
 VERSION AI320733.1 GI:4036715
 KEYWORDS EST.
 SOURCE Neurospora crassa.
 ORGANISM Neurospora crassa
 Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae;
 Neurospora.

REFERENCE 1 (bases 1 to 79)
 AUTHORS Zhu.H., Lai.H., Kupfer.D., Dunlap,J.C. and Roe,B.A.
 TITLE Two Neurospora crassa EST Databases
 JOURNAL Unpublished (1998)
 COMMENT On May 18, 1998 this sequence version replaced gi:3138462.
 Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for genome technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762

Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
 Seq primer: Universal Forward Primer
 High quality sequence stop: 56.
 Location/Qualifiers
 1. .79
 /organism="Neurospora crassa"
 /strain="bd, frq7 A"
 /db_xref="taxon:5141"
 /clone="c9f03nm"
 /clone_lib="Neurospora crassa morning cDNA library"
 /tissue_type="tissue harvested following 22hr growth in dark"
 /note="Vector: pBluescript SK-; Site1: XbaI; Site2: EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:13096,1996. 5' end of cDNA cloned into XbaI site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript;"

BASE COUNT 12 a 12 c 11 g 44 t
 ORIGIN

Query Match 1.3%; Score 39; DB 38; Length 79;
 Best Local Similarity 100.0%; Pred. No. 7.4e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||
 Db 40 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2

RESULT 16
 AI946912 139 bp mRNA EST 17-AUG-1999
 LOCUS bs32h08.y1 Drosophila melanogaster adult testis library Drosophila
 DEFINITION melanogaster cDNA clone bs32h08 5', mRNA sequence.
 ACCESSION AI946912
 VERSION AI946912.1 GI:5737338
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 139)
 Andrews, J., Bouffard, G. and Oliver, B.
 Drosophila melanogaster testis expressed sequence tags
 Unpublished (1999)
 Contact: Brian Oliver
 Laboratory of Cellular and Developmental Biology
 NIDDK, National Institutes of Health
 6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
 Fax: (301) 496 5239
 Email: oliver@helix.nih.gov,
 http://www.niddk.nih.gov/intram/people/boliver.htm
 Tissue isolation and library construction performed at the National
 Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
 http://www.niddk.nih.gov/intram/people/boliver.htm). DNA sequencing
 and analyses performed by National Institutes of Health Intramural
 Sequencing Center (NISC; see http://www.nisc.nih.gov). Chromatogram
 data were analyzed and evaluated for high quality using the ted
 program (Gleeson T and Hillier L, 1991).
 Plate: 32 row: h column: 08
 Seq primer: M13RP1 reverse primer (ABI).
 Location/Qualifiers
 1. .139
 /organism="Drosophila melanogaster"
 /strain="y[*] w[67cl]/Y"
 /db_xref="taxon:7227"
 /clone="bs32h08"
 /clone_lib="Drosophila melanogaster adult testis library"
 /sex="male"
 /dev_stage="1-5 day adult"

FEATURES
 source

/lab_host="SOLR (Stratagene)"
 /note="Organ: testis; Vector: pBluescript SK (Stratagene);
 Site1: EcoR I; Site2: Xho I; Testes dissected from 1-5
 day adult y[*] w[67cl]/Y males raised at 25OC. RNA
 isolated using Trizol (Life Technologies) and a single
 round of Poly(A)+ selection using Oligotex (Qiagen). cDNA
 library constructed using Stratagene ZAP-cDNA syntesis
 kit. Oligo dt-primed, size fractionated -1-6 kb, and
 directionally cloned at EcoRI and XhoI in Uni-ZAP XR.
 Following a single round of amplification pBluescript SK
 phagemids were mass excised. A distribution channel for
 clones is being sought, but not currently available.
 Requests for clones cannot be honored."

BASE COUNT 74 a 30 c 26 g 7 t 2 others
 ORIGIN

Query Match 1.3%; Score 39; DB 46; Length 139;
 Best Local Similarity 100.0%; Pred. No. 7.3e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||
 Db 69 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 107

RESULT 17
 AI540624 161 bp mRNA EST 14-APR-1999
 LOCUS tn88e09.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2176648 3',
 DEFINITION mRNA sequence.
 ACCESSION AI540624
 VERSION AI540624.1 GI:4457997
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 161)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL On May 18, 1998 this sequence version replaced gi:3136848.
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1843 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 116.
 Location/Qualifiers
 1. .161
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2176648"
 /clone_lib="NCI_CGAP_Ut2"
 /tissue_type="moderately-differentiated endometrial
 adenocarcinoma, 3 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site1: SalI;
 Site2: NotI; Cloned unidirectionally. Primer: Oligo df.
 Average insert size 1.85 kb. Life Technologies catalog #:
 11539-012"

BASE COUNT 34 a 10 c 35 g 82 t
 ORIGIN

Query Match 1.3%; Score 39; DB 41; Length 161;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||
 Db 39 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 18
 AA231085 167 bp mRNA EST 26-FEB-1997
 LOCUS mw11ell.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone
 DEFINITION IMAGE:670412 5', mRNA sequence.

ACCESSION AA231085
 VERSION AA231085.1 GI:1853383
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 167)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilsson, R. and
 Waterston, R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:410116
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 155.

FEATURES
 Location/Qualifiers
 1..167
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:670412"
 /clone_lib="Soares mouse 3NME12 5"
 /sex="unknown"
 /tissue_type="fetus"
 /dev_stage="12.5dpc total fetus"
 /lab_host="PH10B"
 /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
 with a modified polylinker; Site.1: Not I; Site.2: Eco RI;
 1st strand cDNA was primed with a Not I - oligo(dT) primer
 [5' GTTACCAATCGAGTGGAGCGCGCCGCTATTTTATTTTATTTT
 3'], on total mouse RNA [provided by Minoru Ko, Wayne
 State Univ.]; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT7T3 vector.
 Library went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 97 a 20 c 18 g 32 t
 ORIGIN

Query Match 1.3%; Score 39; DB 23; Length 167;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||

Db 92 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 130
 |||||

RESULT 19
 AI630474 182 bp mRNA EST 08-MAR-2000
 LOCUS ad1la06.y1 Proliferating Erythroid Cells (LCB:ad library) Homo
 DEFINITION sapiens cDNA clone ad1la06 random, mRNA sequence.

ACCESSION AI630474
 VERSION AI630474.1 GI:4681804
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 182)

REFERENCE
 AUTHORS Gubin, A.N., Njoroge, J.M., Bouffard, G.G. and Miller, J.L.
 TITLE Gene expression in proliferating human erythroid cells
 JOURNAL Genomics 59 (2), 168-177 (1999)
 MEDLINE 99339981

COMMENT On May 18, 1998 this sequence version replaced gi:3136683.
 Contact: Jeffery L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases
 Building 10, Room 9N308, National Institutes of Health, Bethesda,
 MD 20892, USA
 Tel: 301 402 2373
 Fax: 301 402 0101
 Email: jmf7f@nih.gov

The 'ad' library was constructed by Alexander Gubin, Ph.D. in the
 Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
 analyses by National Institutes of Health Intramural Sequencing
 Center (NISC). More information available at:
 http://hembase.niddk.nih.gov
 Plate: 11 row: a column: 06
 Seq primer: M13Rpl reverse primer (ABI).

FEATURES
 Location/Qualifiers
 1..182
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ad1la06"
 /clone_lib="Proliferating Erythroid Cells (LCB:ad
 library)"
 /sex="unknown"
 /tissue_type="blood"
 /cell_type="Erythroid Cells"
 /cell_line="Primary Culture of Peripheral Blood
 Mononuclear Cells"
 /dev_stage="Progenitor; EPO responsive CD71+++"
 /lab_host="DH5alpha"
 /note="Organ: blood; Vector: pCRII; Site.1: EcoRI; Site.2:
 EcoRI; Human peripheral blood mononuclear cells were
 cultured in the presence or absence of erythropoietin.
 CD71+++ cells arising only in erythropoietin-supplemented
 medium were isolated by fluorescence activated cell
 sorting. Those cells demonstrated an average of six
 additional doublings in suspension culture and erythroid
 colony formation in methylcellulose. Suppression
 subtractive hybridization was used to construct the ad
 library (tester-sorted CD71+++ cells, driver-unsorted
 cells cultured without erythropoietin)."

BASE COUNT 88 a 37 c 18 g 39 t
 ORIGIN

Query Match 1.3%; Score 39; DB 42; Length 182;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
 |||||

Db 118 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 156
 |||||

RESULT 20

AW477617 190 bp mRNA EST 24-FEB-2000
DEFINITION 15122 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AW477617
VERSION AW477617.1 GI:7047723
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE 1 (bases 1 to 190)
AUTHORS Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A. and Keele, J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT On Jan 6, 2000 this sequence version replaced gi:6674418.
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCCGAGTCAGCAGC
Plate: 7 row: N column: 9
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. .190
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P1G"
/tissue_type="pooled"
/lab_host="DH108"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

FEATURES
source

BASE COUNT 79 a 53 c 20 g 38 t
ORIGIN
Query Match 1.3%; Score 39; DB 73; Length 190;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 140 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 178
|||||

RESULT 21

AW172002 207 bp mRNA EST 15-NOV-1999
DEFINITION 618047c10.y1 618 - Inbred Tassel cDNA Library Zea mays cDNA, mRNA
sequence.
ACCESSION AW172002
VERSION AW172002.1 GI:6431798
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
REFERENCE 1 (bases 1 to 207)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University

JOURNAL
COMMENT

Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 618047 row: C column: 10.
Location/Qualifiers
1. .207
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="618 - Inbred Tassel cDNA Library"
/tissue_type="tassel"
/dev_stage="tassel length from 0.1 to 2.5 cm"
/lab_host="XLOUR"
/note="Organ: tassel; Vector: pAD-GAL4-2.1 (Hybrizap);
Inbred tassel library from Schmidt lab"

FEATURES
source

BASE COUNT 83 a 29 c 35 g 60 t
ORIGIN
Query Match 1.3%; Score 39; DB 64; Length 207;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 151 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 189
|||||

RESULT 22

AW357288 229 bp mRNA EST 03-FEB-2000
LOCUS 40265 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION AW357288
ACCESSION AW357288
VERSION AW357288.1 GI:6861294
KEYWORDS EST.
SOURCE Bos taurus.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 229)
AUTHORS Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grosse, W.M.,
Bennett, G.A., Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A. and
Keele, J.W.
TITLE Design and use of four pooled tissue normalized cDNA libraries for
EST discovery in cattle
JOURNAL Unpublished (2000)
COMMENT On Apr 7, 1998 this sequence version replaced gi:3034614.
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCCGAGTCAGCAGC
Plate: 18 row: G column: 6
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. .229
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"

FEATURES
source

BASE COUNT 83 a 29 c 35 g 60 t
ORIGIN
Query Match 1.3%; Score 39; DB 64; Length 207;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 151 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 189
|||||

/lab_host="DH108"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, thymus,
semitendonosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
96 a 51 c 49 g 33 t

BASE COUNT
ORIGIN

Query Match 1.3%; Score 39; DB 71; Length 229;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 157 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 195

RESULT 23
AA832425/c
LOCUS AA832425
DEFINITION OC99d12.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1357847 3',
mRNA sequence.
ACCESSION AA832425
VERSION AA832425.1 GI:2905524
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 245)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2286298.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/brp/image/image.html
Insert Length: 992 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 227.

FEATURES
source

Location/Qualifiers
1. .245
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1357847"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH108"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGCGCGCTCATTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
42 a 58 c 64 g 81 t

BASE COUNT

ORIGIN

Query Match 1.3%; Score 39; DB 31; Length 245;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 40 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2

RESULT 24

AW056177
LOCUS AW056177
DEFINITION 60005C01.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.
ACCESSION AW056177
VERSION AW056177.1 GI:5928885
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
1 (bases 1 to 281)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT On Jun 15, 1998 this sequence version replaced gi:3222524.
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660005 row: C column: 01.

FEATURES
source

Location/Qualifiers
1. .281
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOUR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
133 a 37 c 33 g 78 t

BASE COUNT
ORIGIN

Query Match 1.3%; Score 39; DB 63; Length 281;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 236 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 274

RESULT 25

AW479416
LOCUS AW479416
DEFINITION 25195 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AW479416
VERSION AW479416.1 GI:7049522
KEYWORDS EST.
SOURCE Bos taurus.
ORGANISM Bos taurus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 307)
 Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A. and Keele,J.W.

TITLE
 Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle

JOURNAL
 Unpublished (2000)

COMMENT
 On Apr 7, 1998 this sequence version replaced gi:3034581.

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smitht@mail.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACACGCTATGACCAT

BACKWARD: GTTTCACGTCACGACG

Plate: 17 row: P column: 12

Seq primer: ATTTAGGTGACACTATAG.

Location/Qualifiers

1..307

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 4BOV"

/tissue_type="pooled"

/lab_host="DH108"

/note="Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."

BASE COUNT 118 a 65 c 39 g 85 t

ORIGIN

Query Match 1.3%; Score 39; DB 73; Length 307;

Best Local Similarity 100.0%; Pred. No. 7.1e-08;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||

Db 258 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 296

RESULT 26

AA237390

LOCUS

AA237390 318 bp mRNA EST 03-MAR-1997
 mx18b02.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:680523 5'
 similar to gb:M14058 COMPLEMENT C1R COMPONENT PRECURSOR (HUMAN);,
 mRNA sequence.

ACCESSION

AA237390

VERSION

AA237390.1

KEYWORDS

EST.

SOURCE

ORGANISM

house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 318)

REFERENCE

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

On Oct 30, 1996 this sequence version replaced gi:1656971.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:420227

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 267.

Location/Qualifiers

1..318

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_lib="IMAGE:680523"

/clone_lib="Soares mouse NML"

/tissue_type="Liver"

/lab_host="DH108"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5',

TGTTACCAATCTGAAGTGGAGCGCGGGAATCTTTTTTTTTTTT 3'];

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

constructed and normalized by Bento Soares and M.Fatima

Bonaldo."

BASE COUNT 128 a 45 c 68 g 77 t

ORIGIN

Query Match 1.3%; Score 39; DB 23; Length 318;

Best Local Similarity 100.0%; Pred. No. 7.1e-08;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||

Db 275 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 313

RESULT 27

AU082521

LOCUS

DEFINITION

AU082521 Rice panicle shorter than 3cm Oryza sativa cDNA clone

E30614, mRNA sequence.

ACCESSION

AU082521

VERSION

AU082521.1

KEYWORDS

EST.

SOURCE

Oryza sativa.

ORGANISM

Oryza sativa

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

1 (bases 1 to 327)

REFERENCE

AUTHORS

Sasaki,T. and Yamamoto,K.

Rice cDNA from panicle (2000)

TITLE

Unpublished (2000)

JOURNAL

COMMENT

Contact: Takuii Sasaki

National Institute of Agrobiological Resources

Rice Genome Research Program

2-1-2 Kannondai,Tsukuba

Ibaraki,

Japan 305

Tel: 0298-38-7441

Fax: 0298-38-7468

Email: tsasaki@abr.affrc.go.jp

PROJECT = 'RGP'.

E30614_6Z.

FEATURES

source

1..327

/organism="Oryza sativa"

/strain="Nipponbare"

/db_xref="taxon:4530"

/clone="E30614"

/clone_lib="Rice panicle shorter than 3cm"

/dev_stage="shorter than 3cm"		/note="Organ: panicle"	
BASE COUNT	103 a	65 c	82 g
ORIGIN	74 t		
3 others			
Query Match	1.3%;	Score 39;	DB 48; Length 327;
Best Local Similarity	100.0%;	Pred. No. 7.1e-08;	
Matches	39;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	3079	ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117	
Db	272	CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 310	
RESULT 28	AA773444	337 bp	mRNA
LOCUS	ab58d05.s1	Stratagene lung carcinoma 937218	Homo sapiens cDNA clone
DEFINITION	IMAGE:845001	3' similar to contains Alu repetitive element;	mRNA
sequence.			
ACCESSION	AA773444		
VERSION	AA773444.1	GI:2825015	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 337)		
JOURNAL	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,		
COMMENT	Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,		
	Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,		
	Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.		
	WashU-NCI human EST Project		
	Unpublished (1997)		
	On Jan 19, 1998 this sequence version replaced gi:2152978.		
	Contact: Wilson RK		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: est@watson.wustl.edu		
	This clone is available royalty-free through LNL; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	Seq primer: -40m13 fwd. ET from Amersham		
	High quality sequence stop: 276.		
FEATURES	Location/Qualifiers		
source	1. .337		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:845001"		
	/clone_lib="Stratagene lung carcinoma 937218"		
	/tissue_type="lung carcinoma"		
	/cell_line="NCI-H69"		
	/dev_stage="cell line NCI-H69"		
	/lab_host="SOLR (kanamycin resistant)"		
	/note="Organ: lung; Vector: pBluescript SK-; Site:1:		
	EcORI; Site:2: XhoI; Cloned unidirectionally. Primer:		
	Oligo dr. Small cell carcinoma cell line NCI-H69. Average		
	insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor		
	sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'		
	CTCGAGTTTCTTTTCTTTTCTTTT 3'		
BASE COUNT	64 a	87 c	95 g
ORIGIN	91 t		
Query Match	1.3%;	Score 39;	DB 30; Length 337;
Best Local Similarity	100.0%;	Pred. No. 7.1e-08;	
Matches	39;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	3079	ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117	
Db	48	CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 10	
RESULT 30	AA7394641	341 bp	mRNA
LOCUS	sh06a07.y1	Gm-c1016	Glycine max cDNA clone
DEFINITION	Gm-c1016-3949	5' similar to TR:065075	065075 HYPOTHETICAL 30.9 KD
PROTEIN ;			mRNA sequence.
AW394641			
ACCESSION	AW394641		

VERSION
KEYWORDS
SOURCE
ORGANISM

AW394641.1 GI:6913111

EST.

soybean.

Glycine max

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE
AUTHORS

1 (bases 1 to 341)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com

Seq primer: -40RP from Gibco.

Location/Qualifiers

FEATURES
source

1..341
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-3949"
/clone_lib="Gm-cl016"
/tissue_type="immature flowers of field grown plants"
/lab_host="XL10-Gold"
/note="Vector: pBluescript II XR; Site.1: EcoRI; Site.2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Strategene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."

BASE COUNT
ORIGIN

122 a 61 c 60 g 98 t

Query Match

Best Local Similarity 1.3%; Score 39; DB 71; Length 341;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3076 gtccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3114

|||||

Db 303 GTTCATCAAAAAAAAAAAAAAAAAAAAAAAAAA 341

RESULT 31
AI529904

LOCUS

DEFINITION

u183h01.v1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:188904I 5' similar to SW:OATP_RAT P46720 SODIUM-INDEPENDENT ORGANIC ANION TRANSPORTER ;, mRNA sequence.

ACCESSION

AI529904

VERSION

KEYWORDS

AI529904.1 GI:4444039

EST.

SOURCE
ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 376)

Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,

Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,

Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,

Waterston, R. and Wilson, R.

The WashU-NCI Mouse EST Project 1999

Unpublished (1999)

On Jan 19, 1998 this sequence version replaced gi:2286451.

Contact: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

MGI:973365

Seq primer: custom primer used

High quality sequence stop: 361.

Location/Qualifiers

FEATURES
source

1..376
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:188904I"
/clone_lib="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-FL3; Site.1: DraIII (CACATGTGTG); Site.2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTGG], digested and cloned into distinct draIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTCTCTAAAGCTGG and 3' end primer CGACCTGCAGCTGAGACACA."

147 a 61 c 70 g 98 t

BASE COUNT

ORIGIN

Query Match 1.3%; Score 39; DB 40; Length 376;

Best Local Similarity 100.0%; Pred. No. 7.le-08;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117

|||||

Db 309 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 347

RESULT 32

D22652

LOCUS

DEFINITION

sequence.

ACCESSION

D22652

VERSION

KEYWORDS

SOURCE

ORGANISM

EST.

Oryza sativa.

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

REFERENCE

1 (bases 1 to 380)

D22652

RICC0739B Rice callus Oryza sativa cDNA clone C0739_52, mRNA

sequence.

ACCESSION

D22652

VERSION

KEYWORDS

SOURCE

ORGANISM

EST.

Oryza sativa.

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

REFERENCE

1 (bases 1 to 380)

```

AUTHORS      Sasaki,T. and Minobe,Y.
TITLE        Rice cDNA from callus
JOURNAL      Unpublished (1994)
COMMENT      On Nov 22, 1999 this sequence version replaced gi:6462319.
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program
2-1-2 Kannondai, Tsukuba
Ibaraki,
Japan 305
Tel: 0298-38-7441
Fax: 0298-38-7468
Email: tsasaki@abr.affrc.go.jp
PROJECT      "RGP",

FEATURES
source
1. .380
/organism="Oryza sativa"
/strain="cultivar Nipponbare, sub_species Japonica"
/db_xref="taxon:4530"
/clone_lib="Rice callus"
/notes="Vector: pBluescript II SK+; Site_1: Sali; Site_2:
NotI; cDNA prepared from rice callus mRNAs by using
oligo(dT) as a primer and ligating to the Sali-NotI site
of pBluescript II SK+ phagemid."
BASE COUNT   122 a 74 c 82 g 99 t 3 others
ORIGIN
Query Match      1.3%; Score 39; DB 81; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 339 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 377

RESULT 33
AI829162/c
LOCUS      wk76a05.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2421296 3',
DEFINITION mRNA sequence.
ACCESSION  AI829162
VERSION     AI829162.1 GI:5449833
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 424)
AUTHORS     NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      On Jan 19, 1998 this sequence version replaced gi:2285558.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1311 Std Error: 0.00
Seq primer: -40UP from Glibco
High quality sequence stop: 376.
Location/Qualifiers
1. .424
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; lsc
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTCCATCTGAAGTGGAGCGGCATATCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT   128 a 70 c 97 g 132 t
ORIGIN

```

```

/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: Sali;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life technologies catalog #:
11548-013"
BASE COUNT   115 a 87 c 94 g 128 t
ORIGIN
Query Match      1.3%; Score 39; DB 44; Length 424;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 43 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 34
AI050855/c
LOCUS      OY47a05.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:1668944 3',
DEFINITION mRNA sequence.
ACCESSION  AI050855
VERSION     AI050855.1 GI:3307660
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 427)
AUTHORS     NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
JOURNAL      Unpublished (1998)
COMMENT      On Jun 15, 1998 this sequence version replaced gi:3224599.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1392 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 419.
Location/Qualifiers
1. .427
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1668944"
/clone_lib="NCI_CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; lsc
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTCCATCTGAAGTGGAGCGGCATATCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT   128 a 70 c 97 g 132 t
ORIGIN
FEATURES
source
1. .424
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1668944"
/clone_lib="NCI_CGAP_Brn23"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; lsc
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTCCATCTGAAGTGGAGCGGCATATCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT   128 a 70 c 97 g 132 t
ORIGIN

```



```

Query Match      1.3%; Score 39; DB 34; Length 427;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 43 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 35
AA610049/c 444 bp mRNA EST 02-MAR-1998
LOCUS af18h01.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1032049
DEFINITION 3', mRNA sequence.
ACCESSION AA610049
VERSION AA610049.1 GI:2458477
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
TITLE Unpublished (1997)
JOURNAL
COMMENT On Jan 19, 1998 this sequence version replaced gi:2284987.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LInL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 956 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 262.
FEATURES
Location/Qualifiers
source
1..444
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1032049"
/sex="male"
/lab_host="Soares_testis_NHT"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5',
TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT-3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 89 a 103 c 101 g 151 t
ORIGIN
Query Match      1.3%; Score 39; DB 28; Length 444;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 50 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 36
AA731682/c 450 bp mRNA EST 27-JAN-1998
LOCUS nw58e10.s1 NCI_CGAP_GCBI Homo sapiens cDNA clone IMAGE:1250826 3'
DEFINITION similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA731682
VERSION AA731682
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 542 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 363.
FEATURES
Location/Qualifiers
source
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1250826"
/clone_lib="NCI_CGAP_GCBI"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI). Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT-3',
I. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 95 a 85 c 84 g 184 t 2 others
ORIGIN
Query Match      1.3%; Score 39; DB 30; Length 450;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 44 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6

RESULT 37
AW249071
LOCUS AW249071 475 bp mRNA EST 07-JAN-2000
DEFINITION 2820891.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820891 5',
mRNA sequence.
ACCESSION AW249071
VERSION AW249071.1 GI:6592064

```

KEYWORDS SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 475)
AUTHORS NIH-MGC <http://www.ncbi.nlm.nih.gov/MGC/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT On Jul 7, 1999 this sequence version replaced gi:5866355.
Other_ESTs: 2820891.3prime
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: DCTD/DRP cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project Clone Distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center
Trimming: cross_match from University of Washington Genome Center
PHRAP suite. Poly-T identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
<http://www.genome.washington.edu>
Plate: LLCW5 row: G column: 4
High quality sequence stop: 397.
Location/Qualifiers
1..475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2820891"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G).. Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
101 a 135 c 127 g 112 t
BASE COUNT
ORIGIN
Query Match 1.3%; Score 39; DB 70; Length 475;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1800 cctcctgtccaaacctgggcagcagtgagcctgtggtt 1838
|||||
Db 306 CCTCCTGCTCCAACTGGCAGCAGTGGAGCCTGTGTT 344
RESULT 38
AIO18664/c 482 bp mRNA EST 27-AUG-1998
LOCUS ov65e05.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1642208
DEFINITION 3' similar to TR:Q63797 Q63797 PROTEASOME ACTIVATOR RPA28 SUBUNIT
ALPHA.; mRNA sequence.
ACCESSION AIO18664
VERSION AIO18664.1 GI:3232462
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 482)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index
Unpublished (1997)
On May 18, 1998 this sequence version replaced gi:3137113.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 871 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 314.
Location/Qualifiers
1..482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1642208"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'
TGTTACCAATCTGAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
111 a 106 c 111 g 154 t
BASE COUNT
ORIGIN
Query Match 1.3%; Score 39; DB 34; Length 482;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 44 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6
RESULT 39
AA393950 499 bp mRNA EST 12-AUG-1997
LOCUS zt78a10.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:728442
DEFINITION 5' similar to gb:U29007.cdsl AMILORIDE-SENSITIVE SODIUM CHANNEL
ALPHA-SUBUNIT (HUMAN);, mRNA sequence.
ACCESSION AA393950
VERSION AA393950.1 GI:2046919
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 499)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Scheilenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1837 Std Error: 0.00
Seq primer: -28m13 rev2 Et from Amersham
High quality sequence stop: 459.
Location/Qualifiers
1..499
/organism="Homo sapiens"
/db_xref="GDB:5925355"
/db_xref="taxon:9606"
/clone_lib="IMAGE:728442"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 111 a 138 c 127 g 121 t
ORIGIN

Query Match 1.3%; Score 39; DB 25; Length 499;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1800 cctcctgtccaccctggcgagcagtcggcgctgtggtt 1838
|||||
DB 372 CCTCCTGTCCAACTGGCGAGCGAGTGGAGCCTGTGGTT 410
|||||

RESULT 40
AA808304/c 506 bp mRNA EST 07-APR-1998
LOCUS oc4if12.sl NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1352303 3'
DEFINITION similar to TR:Q14012 Q14012 CAM KINASE I.;, mRNA sequence.
ACCESSION AA808304
VERSION AA808304.1 GI:2877710
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 506)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1141 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 199.
Location/Qualifiers
1..506
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1367973"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1352303"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 113 a 113 c 142 g 138 t
ORIGIN

Query Match 1.3%; Score 39; DB 31; Length 506;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
DB 49 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 11
|||||

RESULT 41
AA810286/c 510 bp mRNA EST 25-MAR-1998
LOCUS od14h11.sl NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1367973 3'
DEFINITION similar to SW:IGUP_HUMAN Q06323 INTERPERON GAMMA UP-REGULATED
I-5111 PROTEIN PRECURSOR ;, mRNA sequence.
ACCESSION AA810286
VERSION AA810286.1 GI:2879645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 510)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Dec 18, 1997 this sequence version replaced gi:2340745.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 934 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 227.
Location/Qualifiers
1..510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1367973"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"

```

of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo. "
BASE COUNT      116 a      123 c      117 g      163 t
ORIGIN
Query Match      1.3%; Score 39; DB 37; Length 519;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 3079 ccattcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 42 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4

```

RESULT	43
AT000142	
LOCUS	AT000142 520 bp mRNA EST 01-MAY-1998
DEFINITION	AT000142 Apple young fruit cDNA library Malus x domestica cDNA clone af148, mRNA sequence.

REFERENCE
1 (bases 1 to 520)
AUTHORS
Sung, S.-K., Jeong, D.-H., Nam, J., Kim, S.-H., Kim, S.-R. and An, G.
TITLE
Expressed Sequence Tags of Fruits, Peels, and Carpels and Analyses
of mRNA Expression Levels of the Tagged cDNAs of Fruits from the
Fuji Apple
JOURNAL
Unpublished (1998)
COMMENT
On Apr 7, 1998 this sequence version replaced gi:3036618.

```

BASE COUNT
ORIGIN
168 a      61 c  160 g  129 t    2 others
      pfu:*106;average insert size: 0.85"
      /note="Vector: Lambda vector Unizap II; initial
      /dev_stages="7-10 days after anthesis"
      /clone_lib="Apple young fruit cDNA library"
      /clone="af148"
      /pub_xref="taxon:3750"

```

Query Match	Score 35	DB 47	Length 520
Best Local Similarity	100.00%	Pred. No. 7.1e-08	
Matches 39	Conservative 0	Mismatches 0	Indels 0
Gaps 0			
Qy 3079	ccatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	3117	
Db 479	CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	517	
RESULT 44			
AI397624/c			
LOCUS	AI397624	523 bp	mrna EST
DEFINITION	NCS5G117 Subtracted Conidial Neurospora crassa cDNA clone SC5G117 3' similar to ATP synthase subunit 4, mitochondrial precursor, mRNA sequence.		
			08-FEB-1999

ACCESSION A1397624
VERSION A1397624.1 GI:4240709
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae;
Neurospora.
1 (bases 1 to 523)
Nelson, M.A., Kang, S., Braun, E.L., Crawford, M.E., Dolan, P.L.,
Leonard, P.M., Mitchell, J., Armijo, A.M., Bean, L., Blueyes, E.,
Cushing, T., Errett, A., Fleharty, M., Gorman, M., Judson, K.,
Miller, R., Ortega, J., Pavlova, I., Perea, J., Todisco, S.,
Trujillo, R., Valentine, J., Wells, A., Werner-Washburne, M., Yazzie, S.
and Natvig, D.O.
Expressed sequences from conidial, mycelial, and sexual stages of
Neurospora crassa
Fungal Genet. Biol. 21, 348-363 (1997)
97435549
On Apr 7, 1998 this sequence version replaced gi:3036493.
Department of Biology
University of New Mexico
Castetter Hall, Albuquerque, NM 87131, USA
Tel: 505 277 3411
Fax: 505 277 0304
Email: ngp@biology.unm.edu.
Location/Qualifiers
1..523
/organism="Neurospora crassa"
/strain="74-OR23-IV A (FGSC 2489)"
/db_xref="taxon:5141"
/clone="SC5G11"
/clone_lib="Subtracted Conidial"
/sex="Mating Type A"
/tissue_type="Conidia"
/dev_stage="Germinating conidia"
/lab_host="E. coli"
/note="Vector: pBluescript SK (-); Site_1: EcoRI; Site_2:
XhoI; mRNA isolated from germinating conidia, grown in 1x
Vogel's, 2% sucrose for 4.5 hours. cDNA directionally
cloned into pBluescript SK(-) using the Uni-ZAP XR vector
system (Stratagene, La Jolla, CA). Previously identified
highly expressed clones were subtracted from this
library."

BASE COUNT 125 a 131 c 99 g 168 t
ORIGIN

Query Match 1.3%; Score 39; DB 39; Length 523;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 63 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 25

RESULT 45
AW566146
LOCUS AW566146 523 bp mRNA EST 10-MAR-2000
DEFINITION 660062F05.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.
ACCESSION AW566146
VERSION AW566146.1 GI:7227505
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
1 (bases 1 to 523)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University

JOURNAL Unpublished (1999)
COMMENT On May 7, 1998 this sequence version replaced gi:3118896.
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660062 row: F column: 05.

FEATURES
source
1..523
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."

BASE COUNT 105 a 156 c 131 g 130 t 1 others
ORIGIN

Query Match 1.3%; Score 39; DB 74; Length 523;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 ccatacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3117
|||||
Db 478 CCATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 516

Search completed: September 10, 2000, 23:22:18
Job time: 6096 sec

*****STN Columbus *****

FILE 'HOME' ENTERED AT 14:19:52 ON 23 SEP 2000

=> file medline

COST IN U.S. DOLLARS	ENTRY	SINCE FILE	TOTAL
FULL ESTIMATED COST	SESSION	0.15	0.15

FILE 'MEDLINE' ENTERED AT 14:19:58 ON 23 SEP 2000

FILE LAST UPDATED: 22 SEP 2000 (20000922/UP). FILE COVERS 1960 TO DATE.

MEDLINE has been reloaded to reflect the annual MeSH changes made by the National Library of Medicine for 2000. Enter HELP RLOAD for details.

The OLDMEDLINE file segment now contains data from 1958 through 1965. Enter HELP CONTENT for details.

Left, right, and simultaneous left and right truncation are available in the Basic Index. See HELP SFIELDS for details.

THIS FILE CONTAINS CAS REGISTRY NUMBERS FOR EASY AND ACCURATE SUBSTANCE IDENTIFICATION.

=> s epithelial sodium channel/ab,bi

99491 EPTHELIAL/BI
251712 SODIUM/BI
75167 CHANNEL/BI
5419672 AB/FA
166 EPTHELIAL SODIUM CHANNEL/AB
(EPTHELIAL(W)SODIUM(W)CHANNEL)BI (L)
AB/FA)
99491 EPTHELIAL/BI
251712 SODIUM/BI
75167 CHANNEL/BI
216 EPTHELIAL SODIUM CHANNEL/BI
(EPTHELIAL(W)SODIUM(W)CHANNEL)BI
L1 216 EPTHELIAL SODIUM CHANNEL/AB,BI

=> s transgen?/ab,bi

28596 TRANSGEN?/BI
5419672 AB/FA
21527 TRANSGEN?/AB
(TRANSGEN?/BI (L) AB/FA)
28596 TRANSGEN?/BI

L2 28596 TRANSGEN?/AB,BI

=> s l1 and l2

L3 5 L1 AND L2

=> d l - bib ab

YOU HAVE REQUESTED DATA FROM 5 ANSWERS .
CONTINUE? Y(N)?

L3 ANSWER 1 OF 5 MEDLINE
AN 200023558 MEDLINE
DN 20223558
TI Dysfunction of epithelial sodium transport: from human to mouse.
AU Bonny O; Hummler E
CS Institut de Pharmacologie et de Toxicologie, Universite de Lausanne, Lausanne, Switzerland.
SO KIDNEY INTERNATIONAL, (2000 Apr) 57 (4) 1313-8. Ref: 46
Journal code: KVB. ISSN: 0085-2538.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals
EM 200007
EW 20000702
AB The highly amiloride-sensitive ***epithelial***
sodium
channel (ENaC) is an apical membrane constituent of cells of many salt-absorbing epithelia. In the kidney, the functional relevance of ENaC expression has been well established. ENaC mediates the aldosterone-dependent sodium reabsorption in the distal nephron and is involved in the regulation of blood pressure. Mutations in genes encoding ENaC subunits are causative for two human inherited diseases: Liddle's syndrome, a severe form of hypertension associated with ENaC hyperfunction, and pseudohypoaldosteronism (PHA-1), a salt-wasting syndrome caused by decreased ENaC function. ***Transgenic*** mouse technologies provide a useful tool to study the role of ENaC in vivo. Different mouse lines have been established in which each of the ENaC subunits was affected. The phenotypes observed in these mice demonstrated that each subunit is essential for survival and for regulation of sodium

transport in kidney and colon. Moreover, the alpha subunit plays a specific role in the control of fluid absorption in the airways at birth.
Such mice can now be used to study the role of ENaC in various organs and can serve as models to understand the pathophysiology of these human diseases.

L3 ANSWER 2 OF 5 MEDLINE
AN 2000213442 MEDLINE
DN 20213442
TI Salt-sensitive hypertension in endothelin-B receptor-deficient rats.
AU Garpey C E; Ohuchi T; Williams S C; Richardson J A; Yanagisawa M
CS Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.
SO JOURNAL OF CLINICAL INVESTIGATION, (2000 Apr) 105 (7) 925-33.
Journal code: HS7. ISSN: 0021-9738.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Abbreviated Index Medicus Journals; Priority Journals; Cancer Journals
EM 200007
EW 20000704
AB The role of the endothelin-B receptor (ET(B)) in vascular homeostasis is controversial because the receptor has both pressor and depressor effects in vivo. Spotting lethal (sl) rats carry a naturally occurring deletion in the ET(B) gene that completely abrogates functional receptor expression. Rats homozygous for this mutation die shortly after birth due to congenital distal intestinal aganglionosis. Genetic rescue of ET(B)(sl/sl) rats from this developmental defect using a dopamine-hydroxylase (DBH)-ET(B) ***transgene*** results in ET(B)-deficient adult rats. On a sodium-deficient diet, DBH-ET(B);ET(B)(sl/sl) and DBH-ET(B);ET(B)(+/+) rats both exhibit a normal arterial blood pressure, but on a high-sodium diet, the former are severely hypertensive. We find no difference in plasma renin activity or plasma aldosterone concentration between salt-fed wild-type, DBH-ET(B);ET(B)(+/+) or DBH-ET(B);ET(B)(sl/sl) rats, and acute responses to intravenous L-NAME and indomethacin are similar between DBH-ET(B);ET(B)(sl/sl) and DBH-ET(B);ET(B)(+/+) rats. Irrespective of diet, DBH-ET(B);ET(B)(sl/sl) rats exhibit increased circulating ET-1, and,

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)

on a high-sodium diet, they show increased but incomplete hypotensive responses to acute treatment an ET(A)-antagonist. Normal pressure is restored in salt-fed DBH-ET(B);ET(B)(s/s) rats when the ***epithelial*** ***sodium*** ***channel*** is blocked with amiloride. We conclude that DBH-ET(B);ET(B)(s/s) rats are a novel single-locus genetic model of severe salt-sensitive hypertension. Our results suggest that DBH-ET(B);ET(B)(s/s) rats are hypertensive because they lack the normal tonic inhibition of the renal ***epithelial*** ***sodium*** ***channel***.

L3 ANSWER 3 OF 5 MEDLINE
AN 1999345384 MEDLINE
DN 99345384
TI Implication of ENaC in salt-sensitive hypertension.
AU Hummler E
CS Institut de Pharmacologie et de Toxicologie, Université de Lausanne, Switzerland. ehummle@pop-server.unil.ch
SO JOURNAL OF STEROID BIOCHEMISTRY AND MOLECULAR BIOLOGY, (1999 Apr-Jun) 69 (1-6) 385-90. Ref: 49
Journal code: AX4. ISSN: 0960-0760.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals; Cancer Journals
EM 19991001
EW 19991001
AB Arterial blood pressure is critically dependent on sodium balance. The kidney is the key player in maintaining sodium homeostasis. Aldosterone-dependent epithelial sodium transport in the distal nephron is mediated by the highly selective, amiloride-sensitive ***epithelial*** ***sodium*** ***channel*** (ENaC). Direct evidence that of ENaC participates in blood pressure regulation has come from the molecular analysis of two human genetic diseases, Liddle's syndrome and pseudohypoaldosteronism type 1 (PHA-1). Both, increased sodium reabsorption despite low aldosterone levels in Liddle's patients and decreased sodium reabsorption despite high aldosterone levels in PHA-1 patients, demonstrated that ENaC is an effector for aldosterone action. Gene-targeting and classical ***transgenic*** technology

enable the generation of mouse models for these diseases and the analysis of the involvement of the ***epithelial*** ***sodium*** ***channel*** (ENaC) in the progress of these diseases. A first mouse model using alphaENaC ***transgenic*** knockout mice [alphaENaC(-/-)Tg] mimicked several clinical features of PHA-1, like salt-wasting, metabolic acidosis, high aldosterone levels, growth retardation and increased early mortality. Such mouse models will be necessary in testing the involvement of genetic and/or environmental factors like salt-intake in hypertension.

L3 ANSWER 4 OF 5 MEDLINE
AN 1999170548 MEDLINE
DN 99170548
TI Genetic disorders of membrane transport. V. The ***epithelial*** ***sodium*** ***channel*** and its implication in human diseases.
AU Hummler E; Horisberger J D
CS Institut de Pharmacologie et de Toxicologie, Université de Lausanne, CH-1005 Lausanne, Switzerland.
SO AMERICAN JOURNAL OF PHYSIOLOGY, (1999 Mar) 276 (3 Pt 1) G567-71. Ref: 34
Journal code: 3U8. ISSN: 0002-9513.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals
EM 199906
EW 19990603
AB The epithelial Na⁺ channel (ENaC) controls the rate-limiting step in the process of transepithelial Na⁺ reabsorption in the distal nephron, the distal colon, and the airways. Hereditary salt-losing syndromes have been ascribed to loss of function mutations in the alpha-, beta-, or gamma-ENaC subunit genes, whereas gain of function mutations (located in the COOH terminus of the beta- or gamma-subunit) result in hypertension due to Na⁺ retention (Liddle's syndrome). In mice, gene-targeting experiments have shown that, in addition to the kidney salt-wasting phenotype, ENaC was essential for lung fluid clearance in newborn mice. Disruption of

alpha-subunit resulted in a complete abolition of ENaC-mediated Na⁺ transport, whereas knockout of the beta- or gamma-subunit had only minor effects on fluid clearance in lung. Disruption of each of the three subunits resulted in a salt-wasting syndrome similar to that observed in humans.

L3 ANSWER 5 OF 5 MEDLINE
AN 97471032 MEDLINE
DN 97471032
TI A mouse model for the renal salt-wasting syndrome pseudohypoaldosteronism.
AU Hummler E; Barker P; Talbot C; Wang Q; Verdumo C; Grubb B; Gatz J; Burnier M; Horisberger J D; Beermann F; Boucher R; Rossier B C
CS Institut de Pharmacologie et de Toxicologie de l'Université, Rue du Bugnon 27, CH-1005 Lausanne, Switzerland.
ehummle@pop-server.unil.ch
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1997 Oct 14) 94 (21) 11710-5.
Journal code: PV3. ISSN: 0027-8424.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals; Cancer Journals
EM 199801
EW 19980104
AB Aldosterone-dependent epithelial sodium transport in the distal nephron is mediated by the absorption of sodium through the highly selective, amiloride-sensitive ***epithelial*** ***sodium*** ***channel*** (ENaC) made of three homologous subunits (alpha, beta, and gamma). In human, autosomal recessive mutations of alpha, beta, or gammaENaC subunits cause pseudohypoaldosteronism type 1 (PHA-1), a renal salt-wasting syndrome characterized by severe hypovolemia, high plasma aldosterone, hyponatremia, life-threatening hyperkalemia, and metabolic acidosis. In the mouse, inactivation of alphaENaC results in failure to clear lung liquid at birth and in early neonatal death, preventing the observation of a PHA-1 renal phenotype. ***Transgenic*** alphaENaC driven by a cytomegalovirus promoter in mice [alphaENaC(-/-)Tg] rescued the perinatal lethal pulmonary phenotype and partially restored Na⁺ transport in renal, colonic, and pulmonary

THIS PAGE BLANK (USPTO)

epithelia. At days 5-9, however, alphaENaC(-/-)Tg mice showed clinical features of severe PHA-1 with metabolic acidosis, urinary salt-wasting, growth retardation, and 50% mortality. Adult alphaENaC(-/-)Tg survivors exhibited a compensated PHA-1 with normal acid/base and electrolyte values but 6-fold elevation of plasma aldosterone compared with wild-type littermate controls. We conclude that partial restoration of ENaC-mediated Na⁺ absorption in this ***transgenic*** mouse results in a mouse model for PHA-1.

=> file stnguide

COST IN U.S. DOLLARS	ENTRY	SINCE FILE	TOTAL
	SESSION		
FULL ESTIMATED COST	1.85		2.00

FILE 'STNGUIDE' ENTERED AT 14:21:35 ON 23 SEP 2000
USE IS SUBJECT TO THE TERMS OF YOUR CUSTOMER AGREEMENT
COPYRIGHT (C) 2000 AMERICAN CHEMICAL SOCIETY, JAPAN SCIENCE AND TECHNOLOGY CORPORATION, AND FACHINFORMATIONSZENTRUM KARLSRUHE

FILE CONTAINS CURRENT INFORMATION.
LAST RELOADED: Sep 15, 2000 (20000915/UP).

=> file medline embase biosis inpadoc caplus

COST IN U.S. DOLLARS	ENTRY	SINCE FILE	TOTAL
	SESSION		
FULL ESTIMATED COST	0.00		2.00

FILE 'MEDLINE' ENTERED AT 14:28:05 ON 23 SEP 2000
FILE 'EMBASE' ENTERED AT 14:28:05 ON 23 SEP 2000
COPYRIGHT (C) 2000 Elsevier Science B.V. All rights reserved

FILE 'BIOSIS' ENTERED AT 14:28:05 ON 23 SEP 2000
COPYRIGHT (C) 2000 BIOSIS(R)

FILE 'INPADOC' ENTERED AT 14:28:05 ON 23 SEP 2000
COPYRIGHT (C) 2000 European Patent Office, Vienna (EPO)

FILE 'CAPLUS' ENTERED AT 14:28:05 ON 23 SEP 2000
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE 'HELP USAGE/TERMS' FOR DETAILS
COPYRIGHT (C) 2000 AMERICAN CHEMICAL SOCIETY (ACS)

=> d his

(FILE 'HOME' ENTERED AT 14:19:52 ON 23 SEP 2000)

FILE 'MEDLINE' ENTERED AT 14:19:58 ON 23 SEP 2000
L1 216 S EPITHELIAL SODIUM CHANNEL/AB,BI
L2 28596 S TRANSGEN7/AB,BI
L3 5 S L1 AND L2

FILE 'STNGUIDE' ENTERED AT 14:21:35 ON 23 SEP 2000

FILE 'MEDLINE, EMBASE, BIOSIS, INPADOC, CAPLUS' ENTERED AT 14:28:05 ON 23 SEP 2000

=> s l3

'AB' IS NOT A VALID FIELD CODE

L4 23 L3

=> dup rem l4

PROCESSING COMPLETED FOR L4

L5 10 DUP REM L4 (13 DUPLICATES REMOVED)

=> d l- bib ab

YOU HAVE REQUESTED DATA FROM 10 ANSWERS -
CONTINUE? Y/(N)/Y

L5 ANSWER 1 OF 10 CAPLUS COPYRIGHT 2000 ACS
AN 2000:401983 CAPLUS
DN 133:39099

TI Activated cation channel of the osteoblast as a mechanism of bone anabolism

IN Hruska, Keith
PA Barnes-Jewish Hospital, USA
SO PCT Int. Appl., 19 pp.

CODEN: PLOXDD2

DT Patent
LA English
FAN/CNT 1

PATENT NO.	KIND DATE	APPLICATION NO.
DATE		

PI WO 2000034458 A1 20000615 WO 1999-US28828
19991203

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU,

CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,

IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA,

MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA,
ZW, AM, AZ,

BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT,
BE, CH, CY, DE,

DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BI, CF,

CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
PRAI US 1998-110932 19981205

US 1998-111676 19981210

AB The present invention provides ***transgenic*** non-human mammals

which express stretch-activated cation channel ***transgene*** in

osteoblasts. Such pKBpA/alpha-ENaC animals are generated by transforming the mammals with a ***transgenic*** comprising

an

alpha-subunit ***epithelial*** ***sodium***

channel

(alpha-ENaC) cDNA or its variant inserted into a pKBpA gene and fused to

a promoter such as an osteocalcin promoter.

RE.CNT 4

RE

(1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541 CAPLUS

(2) Canessa, Nature 1993, V361, P467 CAPLUS

(3) Hummler, E, Journal of Steroid Biochemistry and Molecular Biology 1999,

V69(1-6), P385 CAPLUS

(4) Lingueglia, Federation of European Biochemical Societies 1993, V318(1), P95 CAPLUS

L5 ANSWER 2 OF 10 MEDLINE DUPLICATE 1
DN 2000223558 MEDLINE

DN 20223558

TI Dysfunction of epithelial sodium transport: from human to mouse.

AU Bonny O; Hummler E

CS Institut de Pharmacologie et de Toxicologie, Universite de

Lausanne,

Lausanne, Switzerland

SO KIDNEY INTERNATIONAL, (2000 Apr) 57 (4) 1313-8. Ref.

46

Journal code: KVB. ISSN: 0085-2538.

CY United States

DT Journal; Article; (JOURNAL ARTICLE)

General Review; (REVIEW)

(REVIEW, TUTORIAL)

LA English

FS Priority Journals

EW 20000702

AB The highly amiloride-sensitive ***epithelial***

sodium

channel (ENaC) is an apical membrane constituent of cells of many

salt-absorbing epithelia. In the kidney, the functional relevance of

ENaC

BEST AVAILABLE COPY

UNAVAILABLE COPY

THIS PAGE BLANK (USPTO)

expression has been well established. ENaC mediates the aldosterone-dependent sodium reabsorption in the distal nephron and is involved in the regulation of blood pressure. Mutations in genes encoding ENaC subunits are causative for two human inherited diseases: Liddle's syndrome, a severe form of hypertension associated with ENaC hyperfunction, and pseudohypoaldosteronism (PHA-1), a salt-wasting syndrome caused by decreased ENaC function. ***Transgenic*** mouse technologies provide a useful tool to study the role of ENaC in vivo. Different mouse lines have been established in which each of the ENaC subunits was affected. The phenotypes observed in these mice demonstrated that each subunit is essential for survival and for regulation of sodium transport in kidney and colon. Moreover, the alpha subunit plays a specific role in the control of fluid absorption in the airways at birth. Such mice can now be used to study the role of ENaC in various organs and can serve as models to understand the pathophysiology of these human diseases.

L5 ANSWER 3 OF 10 MEDLINE DUPLICATE 2
AN 2000213442 MEDLINE
DN 20213442
TI Salt-sensitive hypertension in endothelin-B receptor-deficient rats.
AU Gartezy C E; Ohuchi T; Williams S C; Richardson J A;
Yanagisawa M
CS Howard Hughes Medical Institute, University of Texas
Southwestern Medical
Center, Dallas, Texas 75390, USA
SO JOURNAL OF CLINICAL INVESTIGATION, (2000 Apr) 105
(7) 925-33
Journal code: HS7. ISSN: 0021-9738.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Abridged Index Medicus Journals; Priority Journals; Cancer
Journals
EW 200007
EW 20000704
AB The role of the endothelin-B receptor (ET(B)) in vascular homeostasis is controversial because the receptor has both pressor and depressor effects in vivo. Spotting lethal (sl) rats carry a naturally occurring deletion in the ET(B) gene that completely abrogates functional receptor expression. Rats homozygous for this mutation die shortly after birth due to

congenital distal intestinal aganglionosis. Genetic rescue of ET(B)(sl/sl) rats from this developmental defect using a dopamine-hydroxylase (DBH)-ET(B) ***transgenic*** results in ET(B)-deficient adult rats. On a sodium-deficient diet, DBH-ET(B);ET(B)(sl/sl) and DBH-ET(B);ET(B)(+/+) rats both exhibit a normal arterial blood pressure, but on a high-sodium diet, the former are severely hypertensive. We find no difference in plasma renin activity or plasma aldosterone concentration between salt-fed wild-type, DBH-ET(B);ET(B)(+/+) or DBH-ET(B);ET(B)(sl/sl) rats, and acute responses to intravenous L-NAME and indomethacin are similar between DBH-ET(B);ET(B)(sl/sl) and DBH-ET(B);ET(B)(+/+) rats. Irrespective of diet, DBH-ET(B);ET(B)(sl/sl) rats exhibit increased circulating ET-1, and, on a high-sodium diet, they show increased but incomplete hypotensive responses to acute treatment an ET(A)-antagonist. Normal pressure is restored in salt-fed DBH-ET(B);ET(B)(sl/sl) rats when the ***epithelial*** ***sodium*** ***channel*** is blocked with amiloride. We conclude that DBH-ET(B);ET(B)(sl/sl) rats are a novel single-locus genetic model of severe salt-sensitive hypertension. Our results suggest that DBH-ET(B);ET(B)(sl/sl) rats are hypertensive because they lack the normal tonic inhibition of the renal ***epithelial*** ***sodium*** ***channel***.

L5 ANSWER 4 OF 10 MEDLINE DUPLICATE 3
AN 1999170548 MEDLINE
DN 99170548
TI Genetic disorders of membrane transport. V. The ***epithelial*** ***sodium*** ***channel*** and its implication in human diseases.
AU Hummler E; Horisberger J D
CS Institut de Pharmacologie et de Toxicologie, Universite de Lausanne,
CH-1005 Lausanne, Switzerland.
SO AMERICAN JOURNAL OF PHYSIOLOGY, (1999 Mar) 276
(3 Pt 1) G567-71. Ref: 34
Journal code: 3U8. ISSN: 0002-9513.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals

EM 199906
EW 19990603
AB The epithelial Na⁺ channel (ENaC) controls the rate-limiting step in the process of transepithelial Na⁺ reabsorption in the distal nephron, the distal colon, and the airways. Hereditary salt-losing syndromes have been ascribed to loss of function mutations in the alpha-, beta-, or gamma-ENaC subunit genes, whereas gain of function mutations (located in the COOH terminus of the beta- or gamma-subunit) result in hypertension due to Na⁺ retention (Liddle's syndrome). In mice, gene-targeting experiments have shown that, in addition to the kidney salt-wasting phenotype, ENaC was essential for lung fluid clearance in newborn mice. Disruption of the alpha-subunit resulted in a complete abolition of ENaC-mediated transport, whereas knockout of the beta- or gamma-subunit had only minor effects on fluid clearance in lung. Disruption of each of the three subunits resulted in a salt-wasting syndrome similar to that observed in humans.

L5 ANSWER 5 OF 10 MEDLINE DUPLICATE 4
AN 1999345384 MEDLINE
DN 99345384
TI Implication of ENaC in salt-sensitive hypertension.
AU Hummler E
CS Institut de Pharmacologie et de Toxicologie, Universite de Lausanne,
Switzerland. ehummler@pop-server.unil.ch
SO JOURNAL OF STEROID BIOCHEMISTRY AND MOLECULAR BIOLOGY, (1999 Apr-Jun) 69
(1-6) 385-90. Ref: 49
Journal code: AX4. ISSN: 0960-0760.
CY ENGLAND; United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LA English
FS Priority Journals; Cancer Journals
EM 199910
EW 19991001
AB Arterial blood pressure is critically dependent on sodium balance. The kidney is the key player in maintaining sodium homeostasis. Aldosterone-dependent epithelial sodium transport in the distal nephron is mediated by the highly selective, amiloride-sensitive ***epithelial*** ***sodium*** ***channel*** (ENaC). Direct evidence that

THIS PAGE BLANK (uspio)

dysfunction of ENaC participates in blood pressure regulation has come from the molecular analysis of two human genetic diseases, Liddle's syndrome and pseudohypoaldosteronism type 1 (PHA-1). Both, increased sodium reabsorption despite low aldosterone levels in Liddle's patients and decreased sodium reabsorption despite high aldosterone levels in PHA-1 patients, demonstrated that ENaC is an effector for aldosterone action. Gene-targeting and classical ***transgenic*** technology enable the generation of mouse models for these diseases and the analysis of the involvement of the ***epithelial*** ***sodium*** ***channel*** (ENaC) in the progress of these diseases. A first mouse model using alphaENaC ***transgenic*** knockout mice [alphaENaC(-/-)Tg] mimicked several clinical features of PHA-1, like salt-wasting, metabolic acidosis, high aldosterone levels, growth retardation and increased early mortality. Such mouse models will be necessary in testing the involvement of genetic and/or environmental factors like salt-intake in hypertension.

L5 ANSWER 6 OF 10 EMBASE COPYRIGHT 2000 ELSEVIER SCI B.V.
AN 1998321597 EMBASE
TI ***Epithelial*** ***sodium*** ***channel*** and its implication in the control of blood pressure.
AU Hummler E.
CS Dr. E. Hummler, Institut Pharmacologie Toxicologie, 27 rue du Bugnon,
CH-1005 Lausanne, Switzerland. ehummler@pop-server.unil.ch
SO Kidney and Blood Pressure Research, (1998) 21/2-4 (253-255).
Refs: 14
ISSN: 1420-4096 CODEN: KBPRFC
CY Switzerland
DT Journal; Conference Article
FS 028 Urology and Nephrology
029 Clinical Biochemistry
LA English

L5 ANSWER 7 OF 10 MEDLINE DUPLICATE 5
AN 97471032 MEDLINE
DN 97471032
TI A mouse model for the renal salt-wasting syndrome pseudohypoaldosteronism.
AU Hummler E; Barker P; Talbot C; Wang Q; Verdumo C; Grubb B; Gatzky J;
Burnier M; Horisberger J D; Beermann F; Boucher R; Rossier B C
CS Institut de Pharmacologie et de Toxicologie de l'Université, Rue du Bugnon

27, CH-1005 Lausanne, Switzerland.
ehummler@pop-server.unil.ch
SO PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1997 Oct 14) 94 (21) 11710-5.
Journal code: PV3 ISSN: 0027-8424.
CY United States
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals; Cancer Journals
EM 199801
EW 19980104
AB Aldosterone-dependent epithelial sodium transport in the distal nephron is mediated by the absorption of sodium through the highly selective, amiloride-sensitive ***epithelial*** ***sodium*** ***channel*** (ENaC) made of three homologous subunits (alpha, beta, and gamma). In human, autosomal recessive mutations of alpha, beta, or gammaENaC subunits cause pseudohypoaldosteronism type 1 (PHA-1), a renal salt-wasting syndrome characterized by severe hypovolemia, high plasma aldosterone, hyponatremia, life-threatening hyperkalemia, and metabolic acidosis. In the mouse, inactivation of alphaENaC results in failure to clear lung liquid at birth and in early neonatal death, preventing the observation of a PHA-1 renal phenotype. ***Transgenic*** expression of alphaENaC driven by a cytomegalovirus promoter in alphaENaC(-/-) knockout mice [alphaENaC(-/-)Tg] rescued the perinatal lethal pulmonary phenotype and partially restored Na⁺ transport in renal, colonic, and pulmonary epithelia. At days 5-9, however, alphaENaC(-/-)Tg mice showed clinical features of severe PHA-1 with metabolic acidosis, urinary salt-wasting, growth retardation, and 50% mortality. Adult alphaENaC(-/-)Tg survivors exhibited a compensated PHA-1 with normal acid/base and electrolyte values but 6-fold elevation of plasma aldosterone compared with wild-type littermate controls. We conclude that partial restoration of ENaC-mediated Na⁺ absorption in this ***transgenic*** mouse results in a mouse model for PHA-1.

L5 ANSWER 8 OF 10 BIOSIS COPYRIGHT 2000 BIOSIS
AN 1997/8766 BIOSIS
DN PREV199799307969
TI Genetic rescue of alpha-ENAC knockout mouse: Establishment of

an animal model for pseudohypoaldosteronism (PHA 1).
AU Hummler, Edith (1); Barker, Pierre; Beermann, Friedrich; Verdumo, Chantal;
Gatzky, John; Boucher, Richard; Rossier, Bernard C.
CS (1) Inst. Pharmacol. Toxicol., Lausanne Switzerland
SO Journal of the American Society of Nephrology, (1996) Vol. 7, No. 9, pp. 1281.
Meeting Info.: 29th Annual Meeting of the American Society of Nephrology
New Orleans, Louisiana, USA November 3-6, 1996
ISSN: 1046-6673.
DT Conference; Abstract; Conference
LA English

L5 ANSWER 9 OF 10 BIOSIS COPYRIGHT 2000 BIOSIS
AN 1995/242043 BIOSIS
DN PREV199598256343
TI Analysis of the ***epithelial*** ***sodium*** ***channel*** (ENaC) in ***transgenic*** mice.
AU Hummler, E. (1); Verdumo, C. (1); Beermann, F.; Rossier, B.
(1)
CS (1) Inst. Pharmacol. Toxicol., Lausanne Switzerland
SO Experientia (Basel), (1995) Vol. 51, No. ABSTR., pp. A7.
Meeting Info.: 27th Annual Meeting of the Swiss Societies for Experimental Biology (USGEB/USBBE) Fribourg, Switzerland March 30-31, 1995
ISSN: 0014-4754.
DT Conference
LA English

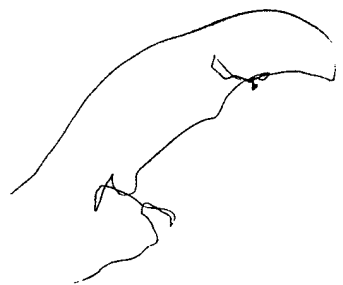
L5 ANSWER 10 OF 10 BIOSIS COPYRIGHT 2000 BIOSIS
AN 1994/239206 BIOSIS
DN PREV199497252206
TI Analysis of the ***epithelial*** ***sodium*** ***channel*** (alpha-rENaC) in ***transgenic*** mice.
AU Hummler, E. (1); Verdumo, C. (1); Beermann, F.; Rossier, B.
(1)
CS (1) Inst. Pharmacol. Toxicol., Lausanne Switzerland
SO Experientia (Basel), (1994) Vol. 50, No. ABSTR., pp. S40.
Meeting Info.: 26th Annual Meeting of the Swiss Societies for Experimental Biology (USGEB/USBBE) Bern, Switzerland March 17-18, 1994
ISSN: 0014-4754.
DT Conference
LA English

=> s 14 and osteoblast#ab/bi

'AB' IS NOT A VALID FIELD CODE
L6 1 LA AND OSTEOBLAST#AB,BI

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)



=> d bib ab

L6 ANSWER 1 OF 1 CAPLUS COPYRIGHT 2000 ACS
AN 2000-401983 CAPLUS

TI Activated cation channel of the ***osteoblast*** as a
mechanism of

bone anabolism

IN Hruska, Keith

PA Barnes-Jewish Hospital, USA

SO PCT Int. Appl., 19 pp.

CODEN: PIXXD2

DT Patent

LA English

FAN CNT 1

PATENT NO. KIND DATE APPLICATION NO.

DATE

PI WO 2000034458 AI 20000615 WO 1999-US28828

19991203

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH,

CN, CR, CU,

CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR,

HU, ID, IL,

IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU,

LV, MA,

MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU,

SD, SE, SG, SI,

SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA,

ZW, AM, AZ,

BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT,

BE, CH, CY, DE,

DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF,

BJ, CF,

CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRAI US 1998-110932 19981205

US 1998-111676 19981210

AB The present invention provides ***transgenic*** non-human

mammals

which express stretch-activated cation channel ***transgene***

in

osteoblasts. Such pKBP/alpha.-rENaC animals are

generated by

transforming the mammals with a ***transgenic*** comprising

an

alpha.-subunit ***epithelial*** ***sodium***

channel

(alpha.-ENaC) cDNA or its variant inserted into a pKBP gene

and fused to

a promoter such as an osteocalcin promoter.

RE CNT 4

RE

(1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541

CAPLUS

(2) Canessa, Nature 1993, V361, P467 CAPLUS

(3) Hummler, E, Journal of Steroid Biochemistry and Molecular

Biology 1999,

V69(1-6), P385 CAPLUS

(4) Lingueglia, Federation of European Biochemical Societies 1993,

V318(1), P95

CAPLUS

=> d his

(FILE 'HOME' ENTERED AT 14:19:52 ON 23 SEP 2000)

FILE 'MEDLINE' ENTERED AT 14:19:58 ON 23 SEP 2000

L1 216 S EPIITHELIAL SODIUM CHANNEL/AB,BI

L2 28596 S TRANSGEN/AB,BI

L3 5 S L1 AND L2

FILE 'STNGUIDE' ENTERED AT 14:21:35 ON 23 SEP 2000

FILE 'MEDLINE, EMBASE, BIOSIS, INPADOC, CAPLUS'

ENTERED AT 14:28:05 ON 23

SEP 2000

L4 23 S L3

L5 10 DUP REM L4 (13 DUPLICATES REMOVED)

L6 1 S L4 AND OSTEOBLAST/AB,BI

=> s l4 and osteocalcin/ab,bi

'AB' IS NOT A VALID FIELD CODE

L7 1 L4 AND OSTEOCALCIN/AB,BI

=> d

L7 ANSWER 1 OF 1 CAPLUS COPYRIGHT 2000 ACS

AN 2000-401983 CAPLUS

DN 133:39099

TI Activated cation channel of the osteoblast as a mechanism of bone

anabolism

IN Hruska, Keith

PA Barnes-Jewish Hospital, USA

SO PCT Int. Appl., 19 pp.

CODEN: PIXXD2

DT Patent

LA English

FAN CNT 1

PATENT NO. KIND DATE APPLICATION NO.

DATE

PI WO 2000034458 AI 20000615 WO 1999-US28828

19991203

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH,

CN, CR, CU,

CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR,

HU, ID, IL,

IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU,

LV, MA,

MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU,

SD, SE, SG, SI,

SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA,

ZW, AM, AZ,

BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT,

BE, CH, CY, DE,

DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF,

BJ, CF,

CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRAI US 1998-110932 19981205

US 1998-111676 19981210

RE CNT 4

RE

(1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541

CAPLUS

(2) Canessa, Nature 1993, V361, P467 CAPLUS

(3) Hummler, E, Journal of Steroid Biochemistry and Molecular

Biology 1999,

V69(1-6), P385 CAPLUS

(4) Lingueglia, Federation of European Biochemical Societies 1993,

V318(1), P95

CAPLUS

=> e hruska keith/au

E1 1 HRUSKA KAREL PROF ING DRSC/AU

E2 3 HRUSKA KATHLEEN S/AU

E3 53 -> HRUSKA KEITH/AU

E4 119 HRUSKA KEITH A/AU

E5 2 HRUSKA KEITH A JR/AU

E6 1 HRUSKA KELTH A/AU

E7 4 HRUSKA KRUNICA/AU

E8 18 HRUSKA L/AU

E9 2 HRUSKA L A/AU

E10 4 HRUSKA L L/AU

E11 5 HRUSKA L S/AU

E12 2 HRUSKA L W/AU

=> s e3-e6

L8 175 ('HRUSKA KEITH'/AU OR 'HRUSKA KEITH A'/AU

OR 'HRUSKA KEITH A

JR'/AU OR 'HRUSKA KELTH A'/AU)

=> s l8 and l1

<-----User Break----->

u

SEARCH ENDED BY USER

If this message appears repeatedly, please notify the Help Desk.

Enter "HELP STN" for information on contacting the nearest STN

Help

Desk by telephone or via SEND in the STNMAIL file.

THIS PAGE BLANK (U0PT0)

=> s 18 and 11

COMMAND INTERRUPTED

'AB' IS NOT A VALID FIELD CODE

L9 5 L8 AND L1

To see the field codes for search terms in an L-number, enter

"DISPLAY

QUER Y" followed by the L-number at an arrow prompt. To see the field

codes for search terms in a saved query, enter "ACTIVATE" and the query name, followed by /Q at an arrow prompt.

=> file medline embase biosis inpadoc caplus

COST IN U.S. DOLLARS SINCE FILE TOTAL

FULL ESTIMATED COST ENTRY SESSION 50.73 52.73

DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)

SINCE FILE TOTAL

CA SUBSCRIBER PRICE ENTRY SESSION -1.11 -1.11

FILE 'MEDLINE' ENTERED AT 14:35:21 ON 23 SEP 2000

FILE 'EMBASE' ENTERED AT 14:35:21 ON 23 SEP 2000

COPYRIGHT (C) 2000 Elsevier Science B.V. All rights reserved.

FILE 'BIOSIS' ENTERED AT 14:35:21 ON 23 SEP 2000

COPYRIGHT (C) 2000 BIOSIS(R)

FILE 'NPADOC' ENTERED AT 14:35:21 ON 23 SEP 2000

COPYRIGHT (C) 2000 European Patent Office, Vienna (EPO)

FILE 'CAPLUS' ENTERED AT 14:35:21 ON 23 SEP 2000

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE 'HELP USAGETERMS' FOR DETAILS.

COPYRIGHT (C) 2000 AMERICAN CHEMICAL SOCIETY (ACS)

=> s 18 and 11

'AB' IS NOT A VALID FIELD CODE

L10 5 L8 AND L1

=> dup rem 110

PROCESSING COMPLETED FOR L10

L11 3 DUP REM L10 (2 DUPLICATES REMOVED)

=> d 1 - bib ab

YOU HAVE REQUESTED DATA FROM 3 ANSWERS -

CONTINUE? Y(N)?

L11 ANSWER 1 OF 3 CAPLUS COPYRIGHT 2000 ACS

AN 2000-401983 CAPLUS

DN 133-39099

TI Activated cation channel of the osteoblast as a mechanism of bone anabolism

IN ***Hruska, Keith***

PA Barnes-Jewish Hospital, USA

SO PCT Int. Appl., 19 pp.

CODEN: PXXXX2

DT Patent

LA English

FAN CNT 1

PATENT NO. KIND DATE APPLICATION NO.

DATE

PI WO 2000034458 A1 20000615 WO 1999-US28828

19991203

W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH,

CN, CR, CU,

CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR,

HU, ID, IL,

IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU,

LV, MA,

MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU,

SD, SE, SG, SI,

SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA,

ZW, AM, AZ,

BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT,

BE, CH, CY, DE,

DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF,

BI, CF,

CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRAI US 1998-110932 19981205

US 1998-111676 19981210

AB The present invention provides transgenic non-human mammals which express

stretch-activated cation channel transgene in osteoblasts. Such

pKBpA.alpha.-rENaC animals are generated by transforming the

mammals with

a transgenic comprising an alpha-subunit ***epithelial***

sodium ***channel*** (alpha.-ENaC) cDNA or its

variant

inserted into a pKBpA gene and fused to a promoter such as an

osteocalcin

promoter.

RE CNT 4

RE

(1) Baker, Molecular and Cellular Biology 1992, V12(12), P5541

CAPLUS

(2) Cantessa, Nature 1993, V361, P467 CAPLUS

(3) Hummler, E. Journal of Steroid Biochemistry and Molecular

Biology 1999,

V69(1-6), P385 CAPLUS

(4) Lingueglia, Federation of European Biochemical Societies 1993,

V318(1), P95

CAPLUS

L11 ANSWER 2 OF 3 CAPLUS COPYRIGHT 2000 ACS

AN 1997:355199 CAPLUS

DN 126-341495

TI Reconstitution of stretch-activated cation channels by expression of the

alpha-subunit of the ***epithelial*** ***sodium***

channel cloned from osteoblasts. [Erratum to document

cited in

CA126:197809]

AU Kizer, Neil; Guo, Xiao-Li; ***Hruska, Keith***

CS Renal Division, Barnes-Jewish Hospital, Washington University

Medical

Center, St. Louis, MO, 63110, USA

SO Proc. Natl. Acad. Sci. U. S. A. (1997), 94(8), 4233

CODEN: PNASAG; ISSN: 0027-8424

PB National Academy of Sciences

DT Journal

LA English

AB On pp. 1013 and 1017 the unit of conductance "picosiemens" was

incorrectly

abbreviated as "psec" and "ps" instead of "pS".

L11 ANSWER 3 OF 3 BIOSIS COPYRIGHT 2000 BIOSIS

DUPLICATE 1

AN 1997:127414 BIOSIS

DN PREV199799419227

TI Reconstitution of stretch-activated cation channels by expression

of the

alpha-subunit of the ***epithelial*** ***sodium***

channel

cloned from osteoblasts.

AU Kizer, Neil; Guo, Xiao-Li; ***Hruska, Keith (1)***

CS (1) Renal Div., Barnes-Jewish Hosp., Washington Univ. Med.

Cent., 216

SO South Kingshighway Blvd., St. Louis, MO 63110 USA

SO Proceedings of the National Academy of Sciences of the United

States of

America, (1997) Vol. 94, No. 3, pp. 1013-1018.

ISSN: 0027-8424.

DT Article

LA English

AB Osteoblasts respond to repetitive strain by activating

stretch-activated,

nonspecific cation channels (SA-CAT) and increasing matrix

protein

production. SA-CAT channels are thought to be responsible for

mechano-transduction in osteoblasts, although the molecular

identity of

the SA-CAT channel has previously been unknown. We have

demonstrated that

both the UMR-106 osteoblast-like cell line and human osteoblasts

in

primary culture express the alpha-subunit of the ***epithelial***

sodium ***channel*** (alpha-ENaC). The ENaC

gene product is

closely related to a class of proteins that confer touch sensitivity to

Caenorhabditis elegans and are referred to as degenerins. A cDNA

THIS PAGE BLANK (USPTO)

clone was obtained of the entire coding region of rat alpha-ENaC (alpha-rENaC). Sequence analysis indicated that the osteoblast clone's sequence was identical to that originally cloned from rat colon. The alpha-rENaC cDNA was cloned into an expression plasmid and transfected into LM(TK-) cells, a null cell for SA-CAT activity. Stable transfectants expressed mRNA and the expected 74-kDa protein corresponding to alpha-rENaC. Reconstitution of alpha-rENaC resulted in the expression of a 24.2 +/- 1.0 psec channel (P-Na:P-K = 1.1 +/- 0.1). The channel is calcium permeable (P-Na:P-Ca = 1.4 +/- 0.1) and highly selective for cations over anions (P-Na:P-Cl mchgr 20). The channel is only active after negative pressure is applied to cell attached patches, cell swelling, or patch excision. These results represent the first heterologous expression of an SA-CAT channel in a mammalian cell system and provide evidence that the ENaC/degenerin family of proteins are capable of mediating both transepithelial sodium transport and are involved in signal transduction by mechano-sensitive cells such as osteoblasts.

=> d his

(FILE 'HOME' ENTERED AT 14:19:52 ON 23 SEP 2000)

FILE 'MEDLINE' ENTERED AT 14:19:58 ON 23 SEP 2000

L1 216 S EPITHELIAL SODIUM CHANNEL/AB,BI
L2 28596 S TRANSGEN?/AB,BI
L3 5 S L1 AND L2

FILE 'STNGUIDE' ENTERED AT 14:21:35 ON 23 SEP 2000

FILE 'MEDLINE, EMBASE, BIOSIS, INPADOC, CAPLUS' ENTERED AT 14:28:05 ON 23 SEP 2000

L4 23 S L3
L5 10 DUP REM L4 (13 DUPLICATES REMOVED)
L6 1 S L4 AND OSTEOBLAST#/AB,BI
L7 1 S L4 AND OSTEOCALCIN/AB,BI
L8 E HRUSKA KEITH/AU
L9 175 S E3-E6
5 S L8 AND L1

FILE 'MEDLINE, EMBASE, BIOSIS, INPADOC, CAPLUS' ENTERED AT 14:35:21 ON 23 SEP 2000

L10 5 S L8 AND L1

L11 3 DUP REM L10 (2 DUPLICATES REMOVED)

=>

--Logging off of STN--

=>

Executing the logoff script...

=> LOG Y

COST IN U.S. DOLLARS SINCE FILE TOTAL

FULL ESTIMATED COST ENTRY SESSION 17.57 70.30

DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)
SINCE FILE TOTAL

CA SUBSCRIBER PRICE ENTRY SESSION -1.11 -2.22

STN INTERNATIONAL LOGOFF AT 14:36:22 ON 23 SEP 2000

THIS PAGE BLANK (USPTO)